U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

HOW CORRUPTION AFFECTS MIGRATION: A GRAVITY MODEL APPROACH

José Pedro Airosa Verónico

Dissertation

Master in Economics

Supervised by

Sandra Maria Tavares da Silva

Maria Isabel Gonçalves da Mota Campos

2018

Acknowledgments

To my parents and brothers, for all the confidence and support they gave me.

To Sofia, for all the love and patience.

To the rest of my family and friends, for helping me through this journey.

To professors Isabel Mota and Sandra Silva, for their kind counselling and precious wisdom.

To all of the above, my gratitude!

Abstract

This research aims at studying the effects of corruption on international migration through a gravity model. By using information that covers migration inflows into 20 Organisation for Economic Co-operation and Development (OECD) countries from the Institute for Employment Research (IAB) dataset, we examine if corruption is a push and/or pull factor for migrants and if this impact varies by skill level and gender. First, we use the data for 2010 with a Poisson regression and then we consider the full period (1980-2010) with a corresponding fixed effects model. We find that, in accordance with previous literature, corruption is a significant push and also a significant pull factor for overall migration. The novelty of this finding is that this result is significant in our panel gravity model, using not only cross-sectional data but also time-series. Our results also show that unlike what the literature predicts, high skilled individuals may not be the most likely to migrate, but rather seem to be the less likely group to be affected by corruption. Adding to that, we find that the effect of corruption on low skilled migrants may be lessened by their inability to migrate. Another finding that is important to highlight is that corruption at the origin country affects different skilled individuals in a different way than corruption at destination. Finally, we find significant gender differences in the effect of corruption. However, if we simply compare overall female and male migration these differences are not significant. In fact, these gender differences remain hidden unless we add the extra skill level layer of analysis. Our results seem to reflect an extra incentive in origin countries to target the female population segment. Furthermore, this discrimination through corruption appears to be less significant in destination countries.

JEL-codes: D73; F22; J16; J24; O1

Key-words: Corruption; International Migration; Gender; Skills; Economic Development.

Resumo

Neste trabalho estudamos os efeitos da corrupção na migração internacional, através de um modelo gravitacional. Usando dados referentes aos imigrantes de 20 países da Organização para a Cooperação e Desenvolvimento Económico (OCDE) do *Institute for Employment Research* (IAB), examinamos se a corrupção é um fator de atração/repulsão para os migrantes e se o impacto varia por género e nível de competências. Primeiro usamos os dados para 2010, através de uma regressão Poisson, em seguida consideramos o período completo (1980-2010) através do modelo de efeitos fixos correspondente. Os resultados, tal como esperado na literatura, apontam no sentido que a corrupção é um fator significativo de atração/repulsão na migração. A novidade desta descoberta é que os resultados são significativos para o nosso modelo gravitacional em painel, usando não só dados para diferentes países, mas também uma série temporal. Os nossos resultados mostram que, contrariamente ao previsto na literatura, os indivíduos com um nível elevado de habilidade podem não ser os mais afetados pela corrupção, sendo que parecem ser o grupo menos afetado por este fenómeno. Acresce a isto, que o efeito da corrupção nos indivíduos com menos competências pode ser diminuído pela sua incapacidade de migrar. Outra conclusão importante de sublinhar é que a corrupção no país de origem e de destino afeta de formas diferentes indivíduos com níveis de competência diferentes. Finalmente, encontramos diferenças significativas do efeito da corrupção entre géneros. Todavia, a simples comparação da migração por género não evidencia diferenças significativas. De facto, estas diferenças apenas se revelam quando consideramos uma dimensão de análise extra, a do nível de competências. Os resultados obtidos parecem refletir um incentivo extra nos países de origem para que as mulheres se tornem alvos da corrupção. Para além disso, esta descriminação através da corrupção parece ser menos significativa nos países de destino.

Códigos-JEL: D73; F22; J16; J24; O1

Palavras-chave: Corrupção; Migrações Internacionais; Género; Competências; Economia do Desenvolvimento.

Contents

List of tables

Chapter 1. Introduction

The study of corruption, through an economic development perspective, is not new. As a matter of fact, since the final decade of the previous century, there have been some studies dedicated to this topic that laid the theoretical foundations of this phenomenon. For example, Mauro (1995) and Shleifer and Vishny (1993) support the existence of a negative relationship between corruption and economic growth that was very influential on the debate of whether corruption could have positive effects on growth.

Corruption is a multivariate concept and can be particularly hard to define (Yusuf, 2012). Beyond the vagueness associated with the inexistence of a clear concept, there seems to be some consensus that it involves the inappropriate use of a public power for a private gain as in Shleifer and Vishny (1993) or Blackburn and Powell (2011) . Corruption also has some peculiar traits that make it a particularly complex subject, namely the imprecision of its measurement. In fact, authors such as Al-Marhubi (2000), Swamy *et al.* (2001) and Aidt (2009) point out that perceived corruption indexes may not match effective corruption. This last trait is not surprising since one of the main goals of corruption, in most countries, is to pass by unnoticed so that it perpetuates itself, either by fear of the law or, at the very least, fear of the consequences of being discovered as mentioned by Shleifer and Vishny (1993). Another trait of corruption these authors mention is its ability to migrate to alternative activities, that is, if there is a focus to eliminate corruption in a certain area, the associated corruption actions can move to another area less monitored. Moreover, corrupt individuals and organizations are more likely to aim at target sectors less able to defend against or ban such behaviour. Tarek and Ahmed (2017), for instance, refer to the shift in public expenditure from sectors like education and health and into less transparent sectors such as highway construction. If corruption is not tackled and indulgence in corruption does not trigger some punitive reaction, other sectors in the economy are also more likely to exhibit corruptive behaviour, as Shleifer and Vishny (1993) argue.

The influence of corruption on topics such as economic growth (*e.g.* Fisman and Svensson, 2007), foreign direct investment (*e.g.* Bénassy-Quéré *et al.*, 2007), inflation (*e.g.* Blackburn and Powell, 2011), or international trade (*e.g.* Anderson and Marcouiller, 2002) has been studied profusely. However, the relation between corruption and migration is still significantly unexplored. Yet, corruption tends to worsen individual working and living conditions for population and, therefore, it might influence migration decisions (Dimant *et al.*, 2013).

In the framework of development economics, migration is also a major topic. Firstly, due to its *raison d'être*, underneath these population movements lie deep social and economic problems, unsolved to the point that people are forced to restart their lives somewhere else, like Ivlevs and King (2017) argue. Secondly, migration strips a country of its human resources, which aggravates the poverty trap as workers and entrepreneurs leave the region diminishing their human resources, as pointed out by Dimant *et al.* (2013).

We thus aim to study the effects of corruption on international migration stocks, particularly by using information from the Institute for Employment Research (IAB) dataset concerning the period 1980-2010 (5 years intervals) and for 20 OECD destination countries by gender, country of origin and educational level. We proceed with a Poisson regression (with robust standard errors), inspired by gravity models of migration, for our data concerning the year of 2010 and then a fixed effects Poisson regression for the full period 1980-2010. We aim at answering the following questions: Is corruption a significant push and/or pull factor for (overall) migrants? Has corruption different impacts on skilled and non-skilled migration? And, is this effect similar between male and female migrants?

The relevance of this work lies on the profound social and economic implications of migration and corruption, as well as the significant gap in the existing literature, which we aim to fill. As a matter of fact, and as Yusuf (2012) mentions, corruption is considered the second most worrying global issue (1st in 10 developing countries) in a global poll, being only behind extreme poverty. The author also mentions the increasing complexity of the migration phenomenon, namely on countries from the OECD. At last, the reasons that influence migration, especially skilled migration, are worth to study, as it may create a larger gap between developed and developing countries (*e.g.*, Acemoglu *et al.* (2002), Rodrik *et al.* (2004)).

This dissertation is structured as follows. Chapter 2 discusses main related concepts and contributions from the literature, systematizing the major findings on corruption and migration research fields. Chapter 3 explains the methodology used and Chapter 4 presents valuable insights and results. Conclusions, main limitations and future research paths are presented in Chapter 5.

Chapter 2. Corruption and migration: main concepts

2.1 Corruption

In this section, we will focus on presenting the roots of the corruption literature and how it has been developed since it emerged, in order to better understand the complexity of this problem. At the same time, it is important to know the various definitions and measures that have been used and are available in the literature. Therefore, we will present various definitions and measures of corruption without forgetting the importance of the relationship between them for a successful and thorough analysis.

2.1.1 Corruption: the roots

As previously mentioned, the study of corruption, through an economic development perspective, is not new. Until recently the discussion about corruption, although significant, lacked the means through which to measure corruption and explore and successfully prove its relationship with *e.g.*, economic growth and development.

Nevertheless, after the end of the cold war, the world started showing some increasing concerns about corruption and first attempts at measuring it appeared. For instance, the Corruption Perception Index (CPI) by Transparency International, or the control of corruption section of the World Governance Indicators (WGI) by the World Bank Group were created halfway through the 1990s (Heywood and Rose, 2013). The development of these measures, as well as a growing interest about corruption and the increasing international pressure to face this issue, lead to the first empirical works (*e.g.* Mauro, 1995) that dealt with corruption and its effects (Jain, 2001).

Shortly afterwards many works studied the influence of corruption on topics such as economic growth (*e.g.* Fisman and Svensson, 2007), foreign direct investment (*e.g.* Bénassy-Quéré *et al.*, 2007), income inequality (*e.g.* Gupta *et al.*, 2002), public debt accumulation (*e.g.* Tarek and Ahmed, 2017), inflation (*e.g.* Blackburn and Powell, 2011), or international trade (*e.g.* Anderson and Marcouiller, 2002), so we can see that corruption has been studied profusely.

An important point that we would like to highlight is that in many of these studies, corruption is used to measure institutional quality or governance capability; usually higher levels of corruption are linked to weaker institutions and lower governance capability. The literature about institutions is responsible for many significant developments and findings about corruption. An example is the review work of Bardhan (1997) on corruption that exposes alternative views of corruption: the "grease the wheels" *vs.* "sand the wheels hypothesis".

Part of the literature supported the hypothesis that corruption might increase countries efficiency, allowing to overcome problems associated with excessive regulations - this is known as the "grease the wheels" hypothesis. On the other hand, and since there are some flaws to the previous argument, for instance, that officials will start purposely delaying the bureaucratic process even more to attract more bribes harming the economy, a different branch of the literature emerged called the "sand the wheels" hypothesis.

Empirical evidence to that point was unable to clarify this issue and one of the reasons why was because there wasn't an indicator to measure corruption. As Jain (2001) claims there were some ambiguous conclusions obtained by the usage of imprecise measures that would have to be subjected to a re-confirmation as better measures of corruption (and governance) arise and develop.

Some authors, like Aggrey (2012), Heywood and Rose (2013) and Tarek and Ahmed (2017), claim there is something like a consensus nowadays or, at least, that most studies support that corruption has overall negative effects on economic performance or on specific indicators, that is, stand for the "sand the wheels" perspective. Nevertheless, there are still few works that argue in favour of the opposing view, or at least partly support it under specific circumstances (Tarek and Ahmed, 2017). A very important part of this puzzle is the concept used to define corruption that will restrain the measurement used, and so it is important to clarify usual concepts and measures of corruption.

2.1.2 Corruption: definitions and measures

A key concept for our work is corruption, which is a multivariate concept and particularly hard to define (Yusuf, 2012). There are many definitions, for instance, Shleifer and Vishny (1993) conceive it as government corruption, defined "as the sale by government officials of government property for personal gains" (*Ibid.*, p. 599). A different definition is given by Blackburn and Powell (2011) "embezzlement of public funds which leads to a loss of resources available to the government for financing its expenditures" (*Ibid.*, p. 225). Dimant *et al.* (2013) defines financial corruption, Tarek and Ahmed (2017) focus on public corruption and Yusuf (2012) stresses political corruption. Jain (2001) also has a different conceptualisation of corruption. The author defines different types of corruption based on the kind of decision that will be influenced by corruption and on what kind of power is held by the decision maker: grand corruption, bureaucratic corruption and legislative corruption. It is clear, that the definition used changes with the purpose of the research.

The instrument used to measure this phenomenon also affects the kind of corruption a study can deal with. As a matter of fact, corruption has some peculiar traits that make it a particularly complex subject, namely the imprecision of its measurement. Jain (2001) also highlights that the need to hide these corrupt actions contribute to blurring their actual impact. Furthermore, corruption spreads across the economy rather than just confining itself. The author claims that if there are opportunities to extract rents, a corrupt regime might be unable to eliminate corruption, and is unlikely to commit to doing so. Finally, the author supports that corruption motivates agents into making decisions and spend funds in channels where collecting bribe is easier; this affects not only quantities spent but also the quality of the investments. Transparency and credibility of punishment are important to avoid the dangers of corruption. This is in line with what we had already mentioned about the literature (*e.g.* Shleifer and Vishny (1993) and Tarek and Ahmed (2017))

So, it is very important to make sure that the kind of corruption we want to measure is in accordance with what is being measured. This is a real obstacle that those who want to study such question must face. There are different kinds of instruments that can be used but not one is without flaws.

Heywood and Rose (2013) divide these instruments as perception-based measurements and non-perceptual approaches. Perception-based measurements such as the CPI, The Bribe Payers Index and WGI (the control of corruption dimension) are the most widely used. The authors argue that there is a difference between the concept and the measurement of corruption, which leads to results that, although reliable, are not necessarily valid in the sense that perceptions of corruption may echo a reality but may be related to issues other than corruption. For instance, authors such as Al-Marhubi (2000), Swamy *et al.* (2001) and Aidt (2009) point out that perceived corruption indexes may not match effective corruption, but even despite highlighting this, sometimes they still have to use these instruments. Some other critics include the fact that the relationship between experience of corruption and its perception may not be linear, and that it usually responds to absolute levels of corruption rather than relative and so, if the level of corruption per person is the same, larger countries will tend to have a higher perceived corruption (Heywood and Rose, 2013). Another argument Heywood and Rose (2013) mention is that these results, given the attention they get, may allow countries to, even those with high levels of corruption, glorify their positive results and dismiss their bad ones, which implies negative consequences for the fight against corruption.

On the other hand, there are the non-perceptual measures. Heywood and Rose (2013) name a few existing instruments such as the analysis of physical infrastructure compared to existing monetary investment, surveys of direct experience of corruption and the rate of criminal convictions of public officials for corruption related crimes. The author also mentions the case of electoral fraud in which regressions and mathematic analysis start being used, although it is very hard to apply such methods into certain aspects of corruption.

Heywood and Rose (2013) consider that many of the previous critiques still apply to these non-perceptual instruments, namely the difficulties of defining corruption and on how these measures can distinguish between different kinds of corruption. Hence, these instruments also fail in allowing a significantly better understanding and organization of this problem. Furthermore, it may not be possible for the instrument to evaluate the seriousness of the act. For instance, an instrument based on the rate of criminal convictions for corruption-related crime, does not discriminate between crimes involving high values and those with low values. Besides that, the lack of convictions may be itself a symptom of a highly corrupt regime and therefore bias this measure. Finally, it may not be possible to scale these measures, to allow comparisons since data may be unavailable due to different reasons such as the costs of developing such a control or lack of records, especially in some developing countries.

In the present work we will use the CPI by Transparency International. This choice is justified in detail in the methodology section. For now, we are focused on presenting the roots of the corruption literature, to understand the complexity of this problem. At the same time, it is important to know how to use this knowledge in order to find the instrument that better fits our goals, and for that we presented various definitions and instruments that have been used and are available in the literature.

In the next section we will describe how previous works studied migration, describing the phenomenon and interpreting other databases. We will also present some theories about migration under an economic development perspective, so that it is easier to establish the relation with corruption, going further and advancing into the works that study both variables.

2.2 Migration

2.2.1 Migration: concepts and measures

As previously mentioned, in the framework of development economics, migration is also a major topic and one that concerns not only developing but also developed countries and that needs international cooperation to be properly dealt with (Docquier *et al.*, 2007).

Firstly, due to its *raison d'être*, underneath these population movements lie deep social and economic problems, unsolved to the point that people are forced to restart their lives somewhere else, like Ivlevs and King (2017) argue. For instance, we can name civil conflicts (Issifou, 2017) or lack of enough attainable income in the origin country and even unemployment as causes for these movements (Vogler and Rotte, 2000). In this line of thought it is also important to distinguish between migrant and refugee.

According to the United Nations, refugees can be defined as "*persons who are outside their country of origin for reasons of feared persecution, conflict, generalized violence, or other circumstances that have seriously disturbed public order and, as a result, require international protection*", whereas for migrants it is stated that although "*there is no formal legal definition of an international migrant, most experts agree that an international migrant is someone who changes his or her country of usual residence, irrespective of the reason for migration or legal status*".[1]

Even though the term refugee is encompassed by the term migrant it is important to distinguish between both terms, since danger and life-threatening circumstances are the main reason for the migration movement of the refugees and not a voluntary choice. Furthermore, they have different legal treatments, for instance, refugees are not penalized for crossing borders without permission and are treated under international laws, whereas other kinds of migrants are treated under national laws.

Secondly, for explaining the relevance of the topic, it is crucial to acknowledge that migration strips a country of its main human resources (phenomenon usually referred as "brain drain"), which aggravates the poverty trap as workers and entrepreneurs leave the region, as pointed out by Dimant *et al.* (2013). However, and despite this being the main perspective in the early stage of the related literature, there are also some studies that point out a positive effect of migration ("brain gain") that may compensate the previous maleficent effect (*e.g.* Mountford (1997) and Mariani (2007)).

[1]http://refugeesmigrants.un.org/definitions, accessed on 12th January 2018.

Several studies have been specifically focused on exploring migration databases, namely Docquier *et al.* (2007), Özden *et al.* (2011), and the extension Docquier *et al.* (2009).

Docquier *et al.* (2007) use an original data set that classifies international migrants by their educational attainment in 1990 and 2000, in order to seek the variables that may be driving brain drain (migration of high skilled workers) from developing countries. The authors claim the lack of such datasets has kept the brain drain discussion on international migrations purely theoretical. Schooling gap and the degree of openness of the sending country are therefore the main focus of this article: it seems that a high skilled migration suggests either a high schooling gap (usually higher in poorer countries) or a high degree of openness (that is affected by country size) but not both at the same time.

According to the authors, overall there is a stronger skilled migration in countries close to the OECD countries and with average levels of schooling that are low (*e.g.* Small Islands of the Pacific and the Caribbean). Proximity also emerges as very important in determining skilled migration flows from Central America. Finally, Sub-Saharan African countries combine a set of different problems such as low level of development and high political instability that promote this kind of migration.

Although there is an increase in the number of migrations by skilled workers from developing countries, emigration rates decreased slightly in this period. This increase associated to skills is explained due to a general rise in educational attainment in these developing countries. Finally, it is interesting to notice that the share of skilled migrants is, in each group, superior to the share of skilled workers among the population (Docquier *et al.*, 2007).

Focusing on the receiving countries, Docquier *et al.* (2007) also give us some interesting insights. In 2000, about two thirds of the total international migration and around sixty percent of the skilled immigrants received by OECD countries were from less developed countries, which is around fifteen p.p. higher than in 1990. Also, in what concerns the recipient countries, in terms of the distribution of migrants into OECD countries, they claim about one fifth of these skilled immigrants live in a member state of the European Union (EU15) and about three quarters of all migrants live in either Australia, Canada and United States. The around five percent skilled migrants left are spread across the other OECD countries.

Özden *et al.* (2011) cover the period from 1960 to 2000, providing additional insights about migration in developing countries. The authors claim that during this period there have been some fundamental changes on the composition of migration, leading to diversified migrant stocks. For example, origin countries are now sending migrants to an increasing number of destination countries. Furthermore, there has been a concentration of migration into developed countries (migration from developing to developed country is the segment that has been increasing at a faster pace, on the contrary, migration from developed to developing countries has been declining) rather than on developing countries that are rather stable in this respect (Özden *et al.*, 2011).

Docquier *et al.* (2009) have developed further the gender dimension of international migration. According to these authors overall female migration has recently exceeded male migration in flows to developed countries due to some determinants such as the rise of the female education attainment or an increased demand for women labour. Özden *et al.* (2011) also cover this issue and support the claim that the composition of immigrants' stocks, in terms of gender, has suffered a significant change.

Docquier *et al.* (2009) mention other reasons why this gender dimension is important to study, for instance women and men may not react with the same intensity to different migration factors (push or pull) such as social networks. Furthermore, there are some barriers that women unlike men might have to face, namely a discriminating work environment that keeps them from reaching their aspirations. This effect may be particularly pronounced in developing countries. Finally, female migration may have a different impact on the economy, for instance, women have been studied to remit more to the origin country than men. There is also an impact at the level of human capital formation due to the brain drain in general, aggravated as Docquier *et al.* (2009) point, since the level of schooling of women is a fundamental ingredient for growth. These are some of the motives that made us include a gender perspective in our analysis and that further explain the relevance of this issue.

The creation of these databases has paved the way for more complex studies as the lack of information and reliable data didn't allow for more thorough analyses. As Vogler and Rotte (2000) claim, insufficient data had long constrained empirical studies about migration flows, especially those from developing countries, so these databases and associated studies represented a step further in the investigation and comprehension of migration.

Based on these databases, there are already studies that link corruption and migration (for example, Poprawe (2015) and Ariu and Squicciarini (2013)). However, they only cover the previous century and so the period after the year 2007, start of the major financial crisis, is not covered at all. In our analysis we use a distinct database which allows us to work with an enlarged time period and this is one of our major contributions to the literature (see chapter 3).

This section served as a way of understanding not only the concepts and measures of migration but also the migration patterns that have been pointed out in the discussion of this topic. In the next subsection we will present studies that are focused on why individuals migrate, more than simply presenting migration movements, and later (section 2.3) we revise studies that discuss corruption as a potential determinant of migration.

2.2.2 Models of Migration

There are many models that, in different ways, discuss the issue of migration. Early works such as Mountford (1997) and Vogler and Rotte (2000) provide some valuable theoretical insights. More recently Beine *et al.* (2016) also provide a theoretical framework for the study of migration, not only supporting some previous theories, but also taking into consideration some new findings made by recent literature.

Vogler and Rotte (2000) support that there is an inverse u-shape relationship between economic development and migration and test the hypothesis that, in early stages of development, economic progress might result in more migration even if there is a decrease on the gap between the origin and the arrival countries. This happens since a higher level of development of the origin country might increase income (or access to funds) or reduce certain costs (*e.g.* information or dislocation costs) and lead to the dissolution of migration restrictions (financial or not) and, *ceteris paribus*, increase migration.

Beine *et al.* (2016) also supports this vision, highlighting the importance of considering a loosening of credit constraints (financial restriction) in the migration decision. Furthermore, the author also highlights the importance of networks that reduces information and uncertainty costs and therefore may also lead to more migration.

However, at a certain point these effects may be diminished since an increase in the development level might reduce differences between countries, lowering migration incentives (*e.g.* income levels become higher). Adding to that, Vogler and Rotte (2000) also point that migration incentives may not depend on absolute income but rather on relative income. Therefore, if an individual has a low income and is surrounded by other individuals in similar economic conditions, there should be an incentive to migrate which is lower than if an individual has a low income and is surrounded by relatively richer individuals.

The existence of a "home preference" discourages individuals from migrating and also leads to lower level of migration than expected (Vogler and Rotte, 2000). Beine *et al.* (2016) suggest that linguistic and cultural proximity between departure and arrival country, for instance, reflect this home preference and can be seen as a non-monetary cost of migration, having an impact on the ability of individuals to adapting to a different country.

Mountford (1997) analyses migration, human capital accumulation and income distribution and how they interact with each other, considering that uncertainty about the success of migration is a key point in the analysis. The author shows that when successful migration is not a certainty, in the sense that there might exist a barrier like immigration regulations that might prevent the individual from migrating and establishing himself successfully, a brain drain can have positive effects on the origin economy, namely by increasing equality and average productivity. Furthermore, even a temporary possibility of migration may have permanent results on the economy. Note that immigration laws are also pointed by Beine *et al.* (2016) as an important factor with influence on migration.

Mountford (1997) contradicts the argument that migration contributes to the lingering of a poverty trap in developing countries due to brain drain to developed countries (Dimant *et al.*, 2013). In fact, since the possibility of migration to a country with higher wages increases the expected return of education, there will be an increase in human capital formation which can more than compensate the negative effect of a brain drain. An important point is that there will be an optimal level of emigration that maximizes these benefits.

These different theoretical insights allow us to go one step ahead of the mere description of migration across countries, enabling a better understanding of individuals, of what may drive them into leaving their countries and what may be the possible consequences of such a movement. Besides these theoretical considerations, it is also important to study this issue empirically to test how well these predictions fit reality. This becomes even more important in cases where different predictions point in opposite directions or create confounding effects. Empirical works like Mayda (2010) and Grogger and Hanson (2011) reached interesting conclusions that connect development issues to migration.

Mayda (2010) studies the determinants of migrant inflows into 14 OECD countries. The author finds that pull factors, namely income opportunities, significantly and positively impact the size of the emigration rate in accordance with what we previously mentioned when referring to Vogler and Rotte (2000). Also in accordance with Vogler and Rotte (2000), distance (that captures the costs of migration) is a significant determinant of emigration rates.

However, push factors, like per worker GDP in the origin country, are rarely negative and when they are negative the size of the effect is insignificant or smaller than for pull factors. This suggests that migration quotas may be more binding than pull effects, increasing the importance of migration policies in host countries. In matter of fact, restricted and exogenous migration quotas may explain this asymmetry between the push and pull factors and, when controlled, that is in the years where immigration policies become less deterring, the effects of both pull and push factors become more significant (Mayda, 2010) .

Grogger and Hanson (2011) create a model of income maximization that makes a cause for positive selection of individuals into migration, since more educated individuals are more likely to emigrate and positive sorting of migrants across destinations, because migrants are more likely to settle in countries with high reward to skill. Migration costs such as linguistic and geographic distance between departure and arrival country, migration networks and colonial heritage are assumed in the model, so this work is in line with the presented literature. Another interesting finding is that post-tax earnings have a stronger correlation with migration than pre-tax earnings. Hence, if corruption is like an implied tax for individuals (Poprawe, 2015), then it might have a similar result, or even a higher negative impact due to its perverse nature. Such relationship needs to be tested in order to better understand this phenomenon. The authors also claim that less educated migrants may be more likely to end up as refugees, therefore countries that give a higher share of visas to refugees, may bias their migration towards less skilled migrants.

Having discussed the topic of migration through an economic development perspective we will now discuss the works that study both corruption and migration and see what has already been figured about this complex issue.

2.3 Migration and Corruption

It is important to review the theoretical mechanisms connecting migration and corruption. Ariu and Squicciarini (2013), Dimant *et al.* (2013) and Ahmad and Arjumand (2016) pay special attention to skilled migration. Dimant *et al.* (2013) say that corruption is a push factor of migration, particularly for skilled migration. They support their results theoretically by claiming that corruption tends to diminish returns to education, damaging the better educated, with less pronounced effects and not statistically robust on average migration. Corruption at home is then a strong incentive for skilled workers to migrate.

Ahmad and Arjumand (2016) make an interesting point, considering that corruption may drive those who do not want to comply with it to either leave their homelands or drives them to become corrupt. The possibility of migration offers skilled and productive human capital a higher income returns on their education investment, and that would be preferable to have it wasted in rent seeking activities. So, besides corruption shifting resources from the productive activities, it may also lead to a significant brain drain. Most importantly, even remittances from the emigrants would not have a significant impact on GDP per capita that could minimize the harm inflicted.

Ariu and Squicciarini (2013) argue that highly skilled workers are mobile, flexible and have a bigger tendency to migrate. These workers tend to leave corrupt countries and search for countries where they gain access to better jobs through skills and merit, and not through nepotism or political affiliation. This would increase out-flows of migrants and decrease in-flows. In their perspective corruption may act more as an obstacle for inflow than a motivation for outflow.

A similar conclusion is reached by Poprawe (2015). The author finds that countries with more corruption attract less immigrants and leads to more emigration, but to all workers and not only high-skilled ones. Corruption seems to be a push factor for migration even when controlling for variables such as distance between countries or GDP per capita. Corruption causes more insecurity and seems to worsen economic conditions, and so individuals migrate to avoid social and economic costs caused by it. Ketterer and Rodríguez-Pose (2015) also supports the idea that less corruption is associated with lower levels of uncertainty and monetary costs. In their study they conclude that certain factors related to regional quality, such as the fight against corruption, attract future residents to certain regions in Europe. They argue that the level of corruption in a region has important financial and non-financial effects.

Other works also tackle this issue but in a more peculiar way. Yusuf (2012) argues that certain aspects of the migration regimes in the UK, and similar ones in other countries, are corruption-friendly in the sense that these countries are the major destination of money laundering by politically exposed persons, privileging these individuals instead of skilled migrant workers who have had stricter laws, constraining them and limiting their mobility and, therefore, also limits the existence of positive externalities in home countries such as know-how transfers. Hence, developed countries might be providing opportunities for certain individuals to not only move illicit funds, but also privileging and offering them additional benefits.

Steinberg (2017) shows that natural resource shocks have a positive brain drain effect in a country, which is especially relevant in countries that are more susceptible to corruption and government inefficiencies. Issifou (2017) also deals with natural resources, showing that migration may reduce natural resource rent seeking and decrease civil conflicts. These authors claim that migration can be as a way of diminishing rent seeking in an economy and through that avoid certain conflicts. This model can be seen as an extension of the model provided by Mariani (2007), where rent seeking is broadly defined and can take the form of corruption, but also of crime, malfunctioning institutions and other activities similar to corruption but that can be linked to it.

The original model provided by Mariani (2007) presented a mechanism through which economies may be influenced by migration. First, the author takes into consideration that the economy can be divided in two parts, namely agents can choose to work in the productive sector of the economy or in the rent seeking sector. Second, the productive sector can be exportable in the sense that they can be used in a different economy (and lead to a brain drain), but the rent seeking sector is not. Finally, if we also introduce a possibility of migration (into a more secure economy, or in other words with a lower rent-seeking sector), then the relative expected return of migration will increase. Also, the proportion of skilled workers that choose to join the rent seeking (that may imply corruption) sector shall decrease and, therefore, the economy of a country will be better off. This work also supports the conclusion by Mountford (1997) that there should be a positive income-maximizing migration rate but differs from the former in the sense that sustains that inefficient allocation of talent is the reason of underdevelopment (and not low levels of education). These results, however, are not robust if we also add endogenous protection: if an individual can use some of its income to protect himself, then the possibility of migration does not guarantee the previous results.

Peng (2009) also extends the model provided by Mariani (2007), by using a framework of heterogeneous ability rather than the homogenous approach of Mariani (2007). Peng (2009) proves that, given the possibility of migration that increases the productive sector attractiveness, such will result not only in a quantitative movement to the productive sector but also in a qualitative movement since individuals with higher talent will dedicate themselves to the productive sector, resulting in a better allocation of talent. In fact, migration increases the comparative payoffs between agents of different productive abilities. This result is both complementary and independent to the one by Mariani (2007).[2]

In accordance with our research questions, it is also important to highlight the segment of the literature that studies the relationship between gender and corruption and explain why it can be important for migration. Swamy *et al.* (2001), Branisa *et al.* (2013) and Jha and Sarangi (2018) are some of the works that tackle this issue.

Swamy *et al.* (2001) claim that women are less likely to participate in or condone bribe taking and, potentially, they may also be more likely to be driven out their country by corruption. This is something we will take into consideration in our analysis (see chapter 4). According to the authors, this claim may be sustained by gender differences in access to corruption networks since women might be less likely to benefit from corrupt activities as they are less likely to be in a position to collect corruption benefits (having lower levels of employment or less experience on how to engage in this kind of activities).

However, Jha and Sarangi (2018) refute the claim that these differences in handling corruption are driven by differences in social status and that this relation might change with equality of gender (the "corruption convergence in gender" hypothesis) and point that these differences will hold even when higher equality is present.

Furthermore, Branisa *et al.* (2013) also show that lower equality and representation in social institutions are related to higher levels of corruption, since as previously mentioned, women might be more averse to corruption and given that they might not have the same opportunities to enrol in the decision making process that may be an additional incentive for women to "vote with their feet" and migrate.

It is relevant to highlight that the literature on corruption and migration is very recent, still being developed. Even though there are some points that seem to be well established in the literature, like the fact that corruption holds a cost for those that migrate, there are still many points that must be studied. Gender differences, education, and corruption and their impact on migration for instance, are relatively unexplored. In matter of fact despite some authors deal with migration and gender (*e.g.* Docquier *et al.*, 2009) or with corruption and gender (for instance, Swamy *et al.* (2001) and Branisa *et al.* (2013)), or even with migration and education (*e.g.* Docquier *et al.*, 2007) and corruption and education

[2] See appendix 6 for a summary table of the literature of corruption and migration.

(*e.g.* Eicher *et al.*, 2009), to the extent of our knowledge, there is no work that addresses these topics simultaneously and that can provide both theoretically and empirical insights.

Chapter 3. Methodology

3.1 Data

In this chapter we will expose the methodology used in this work, focusing on data description and the econometric model. We will start with a brief description of our dependent variable, providing some justifications for the use of the associated measure. We will then do the same for independents variables, with a special highlight on the corruption indicator. Furthermore, we will confront our results with other studies to support the discussion around expected results. As for the model section, we will discuss gravity models and the use of Poisson family models to study migration stocks, both for the year 2010 and for a panel dataset.

3.1.1 Dependent variable

Migration

Our dependent variable is the migration stocks from the IAB dataset (also known as Brücker *et al.* (2013), although we will not use this designation), an expansion of the one used by Docquier *et al.* (2007). This database has information for the period between 1980 and 2010, with five years intervals, and carries information about the inflows for 20 OECD member states. IAB dataset also has the advantage of disaggregating these flows in terms of gender (male or female) and educational attainment namely: *Low Skilled* (primary education or no schooling); *Medium Skilled* (secondary education); *High Skilled* (post-secondary education).

This disaggregation allows us to explore additional possibilities than those already explored in the literature, namely analyse if there are any significant gender differences in migration. At the same time, we can explore questions that have been debated in the literature namely if corruption has different impacts on differently skilled migrants, where in this case, according to the database we are using, skill level is defined by educational attainment.

There is a correspondence between these flows into OECD destination countries and symmetric outflows from 195 origin countries, even though we will only use those that also appear in the "Gravity" database by the *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII), that is, we exclude the data from Monaco, Holy See and Liechtenstein, using 192 different origin countries. We also exclude the migrants with origin "Unknown". By the combination of origin and destination countries we have information on flows for 3820 different country pairs, that will be the "unit" of our analysis, as it is for the case of Poprawe (2015) in a similar study.

In this dataset the definition of migrant relies on the concept of foreign-born rather than of citizenship, although it uses the latter when the first is not available. One reason for this is that country of birth cannot be changed, while in the citizenship case it can vary not only over time, but also from country to country which is not ideal for a time-series and cross-country analysis.

Another important detail is that the definition of migrant also implies individuals with age over 25 years old, in our case this decision is helpful since it allows comparisons with other international migration databases. This criterion also helps excluding individuals who are temporarily studying abroad or that haven't finished their education yet.

Finally, we will use this database for two reasons: first, information obtained in these OECD countries is more reliable than that provided in many destination countries, namely in some developed countries that compose our dataset (Özden *et al.*, 2011). Second, and as previously stated, these countries receive a large part of all international migrants and therefore the size of this database is considerable and relevant for our study (Docquier *et al.*, 2007).

3.1.2 Independent variables

In order to explain the dependent variable, that is the migration stocks, we must take into consideration different variables that have already been used in the literature and see how they behave in our model. These variables can provide answers to the questions we propose to answer so it is important to see how we expect them to perform under our study before we present our results, confronting them with previously established theories.

Corruption

Corruption is the main independent variable in this work, and the one we build our study around. As for the corruption indicator, considering all the above reasoning about the difficulties of defining and accordingly using a specific instrument to measure it, we must account for some additional limitations.

For instance, the period and countries covered by our data (the period between 1980 and 2010) have proven to be an obstacle, since both migration and corruption have a limited number of databases and reliable sources. However, in the case of corruption instruments like this are even more restricted with relatively small time series and a low reach especially within developing countries. Another difficulty arises in knowing what kind of measures should be used since the research question is focused on migrants, being hard to say whether perception or experience of corruption has a bigger weight in migration decisions: first the decision to leave and, then, the decision of where to migrate (if there is such a possibility).

In this work we will use the Corruption Perception Index (CPI) from Transparency International, averaging it for the five-year period. In the cases where this was not possible, namely in the period between 1980 and 1995 we used the estimation of Transparency International that was closest to that date. For 1980 and 1985 we used its estimation for 1980-1985, for 1990 we used the estimation for the 1988-1992 period and for 1995 the estimation for 1993-1996. These estimations have been published in 1996 and expand on the time series but also enlarge the country base of the CPI.

Corruption, according to Transparency International, is defined as "the abuse of entrusted power for private gain. It can be classified as grand, petty and political, depending on the amounts of money lost and the sector where it occurs". [3] The reasons for choosing CPI are: it measures corruption at a national level; it is widely used in the literature (despite the previously mentioned flaws), for instance in Poprawe (2015) and Ahmad and Arjumand (2016); its definition of corruption is close to many works such as Shleifer and Vishny (1993); it is available for a large time period (covering our migration data) and has a wide number of countries listed, that increased from 54 in 1980 to 178 in 2010. We will use as independent variables the CPI score for both the origin and the destination country like in Poprawe (2015), but we will also do some tests with the squared CPI score of the origin and destination countries to see if our results improve. This check for an improvement in the specification has been used in corruption literature about corruption, for instance in Aidt (2009). The CPI for the period covered (it suffered alterations after 2010) can go from 0 to 10 (continuous and not discrete), where a higher value corresponds to a lower perception of corruption. For instance, a change from 5 to 6 implies a reduction in the perceived levels of corruption.

We expect that higher corruption will lead to more migration outflows and lower migration inflows as we described on the literature review (for example, Poprawe (2015) and Ariu and Squicciarini (2013)).

[3] https://www.transparency.org/what-is-corruption#define, accessed on 12th January 2018.

Control Variables

Unemployment rate: The unemployment rate (as a percentage of total labour force), also averaged for each five-year period, is from World Development Indicators. This is an important variable for our study in particular, and for the migration literature in general. White and Buehler (2018), Maria Davidescu *et al.* (2017) and Mayda (2010) are some examples of works that have used this variable to explain migration. Beine *et al.* (2016) suggest the use of this indicator in (gravity) models to study migration since an increase in the unemployment rate may decrease a destination country's ability to attract migrants. This effect may be mitigated by the presence of unemployment benefits in the destination countries. Vogler and Rotte (2000) reach a similar conclusion adding that due to the lack of existence of guarantees that migration will be successful, we can also see the unemployment as connected to the uncertainty of achieving a successful migration.

On the other hand, the level of unemployment in the origin country should have the opposite effect, lack of employment (or uncertainty of employment) may lead individuals to migrate in order to attain some income. Vogler and Rotte (2000) also highlight the role of the diversification of labour markets by the migration of some family members as a way of decreasing the dependence of the family on the labour market of the origin country. This effect is mitigated by the existence of benefits to the unemployed in the origin country.

So, for this indicator we expect that the higher the unemployment rate of the origin country the higher the migration stocks, *ceteris paribus*. On the other hand, a lower unemployment rate of the destination country should lead to higher migration stocks.

Taxes: Tax revenues (% of GDP) are gathered from World Development Indicators. We will use this indicator averaged for the five-year period that composes our data. Grogger and Hanson (2011) and Poprawe (2015) have also used an indicator to take into account the effect of taxes on migration

The rationale behind the use of this instrument is that income is one of the most important determinants for migration. Thus, if the choice is between two different countries with similar characteristics to an individual, choosing the one with the lower taxes (higher available income) is the most obvious choice, although we are aware that this may not be the best indicator to measure this variable. Nevertheless, we use this indicator since it covers a longer period of time (although for some countries the time-series is shorter) and has a wider scale of countries, being broader than other databases used to measure taxation level.

We expect that as taxation level in the destination country increases (decrease in the origin country) there will be a decrease in migration, *ceteris paribus*. However, this effect may be changed since high taxes level also implies some social security. Namely many of the destination countries have some kind of social income that they provide to their citizens.

Inflation: We use a five-year averaged inflation rate, measured by the consumer price index (annual %) from the World Development Indicators. We use the five-year averaged inflation ration. Inflation serves as a measure of macroeconomic instability and, at the same time, it also influences the real income of individuals, inducing their preferences to migrate. According to the literature we expect that more inflation will lead to less migration (Poprawe, 2015) in both the destination (mostly due to macroeconomic instability) and origin countries (loss of real income).

The CEPII provides a few databases that have been widely used by many studies in the trade literature (*e.g.* Burger *et al.*, 2009). Recently many studies in migration have also used the variables from their databases (*e.g.* Poprawe, 2015).

From the Gravity and Geo CEPII databases we use:

Gross Domestic Product (GDP) *per capita* *(*current US $): We use the five-year averaged natural logarithm of the GDP *per capita* in both the destination and origin country. According to the literature, a higher GDP per capita is expected to reflect a higher income at a given country, so we expect that there are more inward migrants and less outward migrants at countries with a higher income (Beine *et al.*, 2016). This effect may be lessened by the existence of credit constraints or by the existence of significant costs caused by home preferences leading to positive coefficients for the GPD per capita at the home country (Poprawe, 2015).

Population: We used the five-year averaged natural logarithm of the total population (in millions) in both the destination and origin country. This indicator is the mass variable of our gravity model and an increase in the population of either destination or origin country will lead to an increase in migration (Poprawe, 2015).

Contiguity: A dummy variable, with value 1 if both countries are contiguous and 0 otherwise. In other words, this variable assumes the value 1 if destination and origin country share a common border. Migration should be higher (Poprawe, 2015) if both countries are contiguous as it reflects proximity (both physical and cultural), however it has also appeared as non-significant in some other studies Mayda (2010).

Common (official primary) language: A dummy variable, with value 1 if both countries have the same official primary language and 0 if they do not. According to the literature (Beine *et al.*, 2016) if countries share the same official primary language then, *ceteris paribus*, migration stocks should be higher. Not knowing the language may prove to be a hard barrier to overcome and that may be a deterrent to migration.

Common religion: This variable varies between 0 and 1. If countries share the same religion with no fractionalization then this variable will equal 1, if there is no common religion despite fractionalization then this will equal 0. Sharing a common religion should increase migration, *ceteris paribus,* reflecting cultural proximity (Beine *et al.*, 2016). However, due to the homogeneity that exists in destination countries this result might be biased so we should be careful when using it.

Distance (population weighted per kilometre): The weighted distance is based on bilateral distances between the major cities of those two countries, with inter-city distances being weighted by the percentage of the city in the overall country's population, according to the formula:

$$d_{ij}=\left(\sum_{k\in i}(pop_k/pop_i)\sum_{l\in j}(pop_l/pop_j)d_{kl}^{\theta}\right)^{1/\theta} \quad (3.1)$$

where pop_k is the population of agglomeration k belonging to country i, and θ measures the sensitivity of migration to bilateral distance d_{kl}. In this case, θ is set equal to -1 and it has the property of constant elasticity of substitution.[4]

We use the natural logarithm of this distance between both countries. According to the literature as the distance increases, migration between two countries should decrease, *ceteris paribus,* reflecting increased migration costs (Vogler and Rotte, 2000).

3.1.3 Summary statistics

Table 1 presents some summary statistics for our database (for 2010). We can see that for our dependent variables there is a very high standard deviation, even above the mean values. In our sample, for the year 2010, we can see that there are more female migrants abroad than male migrants, with the only exception being the medium skilled segment. As for skill level, high skilled migrants are the segment that represents the most individuals. The dummy variables show the percentage of the country pairs with a given attribute, for example

[4] We also experimented our model with the distance variable where the θ value takes the value of 1, but since the overall results do not change much we chose θ=-1 because it is the usual coefficient the literature uses and because of the additional CES property.

less than three per cent of our country pairs have a colonial relationship between the departure and arrival country. Another important point is that for variables like corruption, inflation and other economic variables, the mean for the destination countries presents a better environment for the destination countries than for origin countries.

Table 1- Summary statistics

Variable	Country Pairs	Mean	Standard Deviation	Minimum	Maximum
Overall_Migrants	3820	19163.81	167997.1	0	9234340
Low Skill_Migrants	3820	6095.184	89812.95	0	5292107
Medium Skill_Migrants	3820	5422.156	48003.84	0	2626342
High Skill_Migrants	3820	7646.473	46142.5	0	1315891
Male_Migrants	3820	9490.752	89702.61	0	5044610
Male_Low Skill_Migrants	3820	2982.869	48779.89	0	2900516
Male_Medium Skill_Migrants	3820	2792.36	26401.36	0	1468257
Male_High Skill_Migrants	3820	3715.523	22655.19	0	675837
Female_ Migrants	3820	9673.06	78923.64	0	4189730
Female_Low Skill_Migrants	3820	3112.315	41158.97	0	2391591
Female_Medium Skill_Migrants	3820	2629.795	21976.12	0	1158085
Female_High Skill_Migrants	3820	3930.95	23861.66	0	672968
CPI_Destination	3820	7.965	1.274	4.2	9.4
CPI_Origin	3580	4	2.062	1.15	9.4
Colonial_Relationship	3820	0.026	0.159	0	1
Common_Language	3820	0.142	0.349	0	1
Common_Religion	3780	0.183	0.238	0	0.969
Contiguity	3820	0.015	0.12	0	1
Distance	3800	8.662	0.836	4.952	9.88
GDPPC_Destination	3820	10.666	0.467	9.273	11.527
GDPPC_Origin	3720	8.358	1.542	5.218	11.527
Inflation_Destination	3820	2.116	0.665	0.887	3.6
Inflation_Origin	3540	54.193	637	-0.077	8503.581
Unemp_Destination	3820	6.688	2.31	3.25	12.467
Unemp_Origin	3120	8.678	6.641	0.3	35.5
Pop_Destination	3820	2.637	1.368	-0.715	5.717
Pop_Origin	3800	1.705	2.154	-4.627	7.189
Tax_Destination	3629	21.007	6.729	9.639	33.259
Tax_Origin	2821	17.278	7.438	1.217	48.838

Source: Own computation.

There are other characteristics of migration that are also reflected in our data. For instance, in Table 2 we can see that origin countries are now sending migrants to an increasing number of destination countries as Özden *et al.* (2011) claimed, and that this trend is transversal for female and male migrants. Given our objectives we will not analyse this in excessive detail, even though we could do a more thorough description of these characteristics.

Table 2- Migration stocks between countries

Year	Total Migrants		Female Migrants	
	Country Pairs with 0 flows	% of Country Pairs with 0 flows	Country Pairs with 0 flows	% of Country Pairs with 0 flows
1980	1749	45.79	1881	49.24
1985	1178	30.84	1319	34.53
1990	1025	26.83	1195	31.28
1995	708	18.53	888	23.25
2000	574	15.03	738	19.32
2005	526	13.77	684	17.91
2010	509	13.32	658	17.23

Source: Own computation

After introducing the variables, we will now present the correlation matrix in order to have a first view of the interaction between each pair of variables and to verify if there is any irregularity or strong correlation we should be careful about in our model (Table 3).

According to our results, we have relatively small correlation values between our dependent and independent variables. As for CPI scores the correlation is always below 0.05, with a positive sign for the CPI of the origin with most of the independent variables and a negative one for the CPI of the destination country.

As for the correlation between independent variables, there are some cases that require our attention. The first one is the very high correlation coefficient between GDP *per capita* and the CPI score of the origin country that is over 0.81. However, since this variable is very important for our model we cannot simply overlook it and furthermore this relationship does not go against what we expected. In fact, on one hand high per capita income countries such as Sweden, Denmark and others present low levels of perceived corruption and, on the other hand, lower income economies like Niger or Cambodia present high levels of perceived corruption. Other high correlation coefficients are the ones between the CPI score and the unemployment rate of the destination country (0.67), and GDP *per capita* and the unemployment rate of the destination country (0.61). Again, these results are not surprising given the close link between employment and GDP per capita. Furthermore, this proximity may be strengthened due to the small amount of observations.

Table 3 – Correlation Matrix

	Overall Migrants	CPI Destination	CPI Origin	Colonial Relationship	Common Language	Common Religion	Contiguity	Distance	GDPPC Destination	GDPPC Origin	Inflation Destination	Inflation Origin	Unemp. Destination	Unemp. Origin	Pop. Destination	Pop. Origin	Tax. Destination	Tax. Origin
Overall Migrants	1.000																	
CPI Destination	-0.033	1.000																
CPI Origin	0.001	-0.005	1.000															
Colonial Relationship	0.057 ***	-0.039 *	-0.038 *	1.000														
Common Language	0.065 ***	0.056 ***	0.060 ***	0.263 ***	1.000													
Common Religion	0.010	-0.040 *	0.183 ***	-0.045 **	0.130 ***	1.000												
Contiguity	0.181 ***	-0.017	0.153 ***	-0.022	0.160 ***	0.169 ***	1.000											
Distance	-0.051 **	0.072 ***	-0.224 ***	0.010	0.049 **	-0.096 ***	-0.382 ***	1.000										
GDPPC Destination	0.004	0.499 ***	-0.003	-0.016	0.004	-0.025	0.010	-0.198 ***	1.000									
GDPPC Origin	0.030	-0.001	0.814 ***	-0.063 ***	-0.020	0.203 ***	0.172 ***	-0.318 ***	-0.003	1.000								
Inflation Destination	0.006	-0.336 ***	0.002	0.042 **	0.028	-0.125 ***	-0.043 **	0.244 ***	-0.531 ***	0.001	1.000							
Inflation Origin	-0.032	0.000	-0.537 ***	0.021	-0.004	-0.142 ***	-0.123 ***	0.203 ***	0.001	-0.543 ***	-0.001	1.000						
Unemp Destination	0.015	-0.671 ***	0.003	0.001	-0.049 **	0.079 ***	0.014	-0.051 **	-0.610 ***	0.002	0.140 ***	-0.000	1.000					
Unemp Origin	-0.030	0.002	-0.122 ***	0.004	0.057 ***	0.012	-0.019	-0.047 **	0.002	-0.097 ***	-0.000	0.066 ***	-0.003	1.000				
Pop Destination	0.171 ***	-0.304 ***	0.002	0.143 ***	0.091 ***	-0.047 **	0.019	0.085 ***	-0.340 ***	0.001	0.039 *	-0.000	0.415 ***	-0.001	1.000			
Pop Origin	0.111 ***	0.002	-0.185 ***	-0.053 **	-0.074 ***	-0.159 ***	0.058 ***	0.010	0.002	-0.132 ***	-0.000	0.016	-0.003	-0.169 ***	-0.006	1.000		
Tax Destination	-0.123 ***	0.314 ***	-0.002	0.062 ***	-0.088 ***	-0.089 ***	-0.037 *	-0.021	0.178 ***	-0.000	0.225 ***	-0.000	-0.411 ***	0.001	-0.561 ***	0.003	1.000	
Tax Origin	-0.024	-0.002	0.311 ***	0.024	0.102 ***	0.088 ***	0.027	-0.140 ***	-0.001	0.294 ***	-0.002	-0.176 ***	0.003	0.280 ***	0.004	-0.352 ***	-0.007	1.000

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Source: Own computation.

3.2 The model

3.2.1 Gravity models

As we already stated, in this work we take influence from the Gravity models that are usually used in the international trade literature (*e.g.* Burger *et al.*, 2009) but that have recently been adopted in works about migration (*e.g.* Beine *et al.*, 2016), even though there has been an ongoing discussion on the specification of these models as we will show.

Gravity Models rely on Newton's law of universal gravitation. Accordingly, the gravitational attraction between two countries is proportional to the product of the masses of the countries, usually the Gross Domestic Product (common in trade related studies) or the size of the Population (common in migration studies), and inversely proportional to the geographical distance as presented by Burger *et al.* (2009) in a basic model:

$$I_{ij}=K\frac{M_i^{\beta_1}M_j^{\beta_2}}{d_{ij}^{\beta_3}} \tag{3.1}$$

where I_{ij} is the interaction intensity of our dependent variable, so in our case that would be migration stocks from one country to another, M is the mass of the countries measured by its population, d_{ij} is the distance between two country and betas are parameters to measure the strength of these effects. The model can be easily adapted in order to have variables such as common language or past colonial relationship, or even the CPI score for both origin and destination country.

In earlier stages of the literature, this model was usually formalized through a log-normal specification. This specification however has a couple of disadvantages that started being pointed out in the literature (*e.g.* Burger *et al.*, 2009)., such as the bias that is generated through the logarithmic transformation; the assumption that all error terms have equal variance (homoskedasticity) that did not hold and which puts in cause the efficiency and consistency of the model (Beine *et al.*, 2016); the sensitivity of the model to zero valued flows is also a significant limitation of this specification since there is no way of properly dealing with this flaw in a log-normal scenario as either dropping the zeros or giving these zeros a value of 1 leads to inconsistent estimators (Santos Silva and Tenreyro, 2006). Finally, an ordinary least squares (OLS) model also has the shortcoming of being able to predict negative events when they should take the value of 0 or above (Wooldridge, 2010).

Given these fragilities researchers have turned into different models to find a way of solving this dilemma. The literature has turned into models from the Poisson family, for instance Poprawe (2015) uses a negative binomial model and does a robustness check with a Poisson pseudo-maximum-likelihood (PPML) model. Other examples include the zero inflated variations of the Poisson and negative binomial models such as the one used by White and Buehler (2018).

The Poisson family models are usually used to model count data, that is a model where the dependent variable is a discrete non-negative integer (Hilbe, 2014). In the case of migrations, the number of people that migrate (in our definition foreign born individuals, for most cases) is clearly such a dependent variable.

According to Greene (2011), the primary equation of the model is:

$$\text{Prob}(Y = y_i \mid x_i) = \frac{e^{-\lambda_i}\lambda_1^{y_i}}{y_i!}\cdot, \; y_i = 0, 1, 2, \ldots \tag{3.2}$$

where:

$$\lambda_i = e^{\boldsymbol{x}_i'\boldsymbol{\beta}} \tag{3.3}$$

or in an equivalent way (although we will work with the previous functional form):

$$\ln(\lambda_i) = \boldsymbol{x}_i'\boldsymbol{\beta} \tag{3.4}$$

Additionally, the Poisson model has the restrictions:

$$E[yi \mid xi] = \lambda_i = e^{\boldsymbol{x}_i'\boldsymbol{\beta}} \tag{3.5}$$

$$Var[yi \mid xi] = \lambda_i = e^{\boldsymbol{x}_i'\boldsymbol{\beta}} \tag{3.6}$$

and therefore:

$$E[yi \mid xi] = Var[yi \mid xi] = \lambda_i = e^{\boldsymbol{x}_i'\boldsymbol{\beta}} \tag{3.7}$$

and then the model is solved by a maximum likelihood estimation as shown for instance in Cameron and Trivedi (2009).

The restriction imposed in (3.5) is also known as the Poisson variance assumption and if it holds then the data is equidispersed, however this assumption does not usually hold. In our case our dependent variables are clearly overdispersed, in fact a simple alpha test (through a likelihood ratio test of the Poisson model *vs.* the "negative binomial 2" (NB2) model) asserting that the probability that we would observe these data conditional on α=0 is virtually zero, favouring the overdispersion hypothesis ($\alpha > 0$) and clearly rejecting equidispersion in all the sub-samples (by skill and by gender) of our sample.[5]

In order to effectively deal with overdispersion there are a couple of different adaptations to the model that we make, and that have been made in the literature.

[5] We provide the alpha values that we found in the table of the NB2 regression in appendix 2 (and 3).

The first alternative involves the use of Poisson with robust standard errors as suggested by Santos Silva and Tenreyro (2006), that is the estimation is made by pseudo-maximum-likelihood (or quasi-maximum-likelihood). These authors and others such as Cameron and Trivedi (2009) claim that this method provides consistent estimates even if the count is not actually Poisson distributed. Furthermore, it is consistent even if the dependent variable is not a count and that is the reason why it can be used to predict, for example, the volume of exports in trade literature.

The PPML unlike the regular Poisson model only requires that the conditional mean function specified in equation (3.5) and does not require the equation (3.6) to hold to maintain relative efficiency as is discussed in Wooldridge (2010). This author claims that, due to its robustness, this alternative has an advantage over some other alternatives such as the "Negative Binomial 1" (NB1) model where the variance assumption is relaxed to (according to Hilbe (2014)):

$$\text{Var}\ [y_i \mid x_i] = \lambda_i(1+\alpha) \tag{3.8}$$

If we wanted to estimate conditional probabilities then a different model such as NB2, where the variance assumption is relaxed to (according to Hilbe (2014)):

$$\text{Var}\ [y_i \mid x_i] = \lambda_i(1+\alpha\,\lambda_i) \tag{3.9}$$

might also provide a more flexible alternative, however we are more interested in finding the effects of our independent variables on migration and not as much in finding the probability of having a specific number of migrants (and therefore we did not specify any density function for the negative binomial models in this section), so the PPML model seems to be a good fit.

3.2.2 Panel data models

After estimating the PPML model for 2010, we will then use the full dataset, including information for the period 1980-2010 (although the benefit of additional information is diminished by the lack of data on CPI for some countries).

For this, we will use a two-way (time and country pair) fixed effects Poisson model as specified by Fernández-Val and Weidner (2016):

$$\text{E}\ [y_{it} \mid x_{it}] = \lambda_{it} = e^{x_{it}'\boldsymbol{\beta} + \delta_t + \theta_i} \tag{3.10}$$

where we add the term δ_t to account for time-fixed fixed effects and θ_i to account for country pair fixed effects. Notice that the inclusion of country pair fixed effects allows for the model to account for multilateral resistance to migration, that is some disturbance in migration caused by the existence of alternative and attractive destinations (Beine *et al.*, 2016).

The fixed effects Poisson Regression (with cluster-robust standard errors), similarly to PPML, is able to consistently estimate the conditional mean parameters and therefore fits to our purposes. As shown in Wooldridge (1999) (and also in Wooldridge, 2010), this model is completely robust to distributional misspecification and serial correlation. Notice that the random effects model, unlike the fixed effects model, has no known robustness properties and therefore would need the Poisson distribution assumption, serial independence and some other assumptions to hold. For instance, the fixed effects model, unlike the random effects, allows for unrestricted heterogeneity across individuals. Since we believe that this heterogeneity may have a significant influence on our results, country pairs fixed effect cannot be neglected as stated by Baltagi *et al.* (2014). Therefore, a random effects model does not provide any additional useful advantages when compared to the fixed effects model. Furthermore, a Hausman test would not be useful to decide between these two models since we are using robust standard errors,[6] like it is to decide between for the negative binomial fixed effects model (NBFE), that is chosen *vs.* its random effects counterpart. Last but not least, in the literature there is a clear preference for fixed effects over random effects models (*e.g.* Poprawe, 2015)

The negative binomial fixed effects model on the other hand is also not a solution to this problem and has been contested in the literature. For instance, Allison and Waterman (2002) show that this is not a true fixed effects estimator, as with this model it is possible to calculate estimates for time invariant variables, and so the authors suggest using different approaches. This happens since, as the authors claim, the model is based on a regression decomposition of the overdispersion parameter rather than a decomposition of the mean.

Guimarães (2008) suggests a test to see if the NBFE can be appropriate despite the previous arguments, but we cannot use it since for this test it is recommended at least 20 time periods for 1000 individuals. So, our data does not fit the requirements to perform this test and therefore we will not use this model.

[6] If we were not using robust standard errors, the results would point to fixed-effects, but this is meaningless given the assumption failures, so we will not present this test.

Chapter 4. The influence of corruption on migration stocks: an empirical assessment

In this chapter we implement an econometric procedure in order to analyse the impact of corruption on migration stocks, after controlling for other determinants of migration. We will start by implementing the Poisson model with robust standard effects (PPML) for 2010. We will then check our results against the fixed effects version of the Poisson model that also works as a robustness test in other similar works (Poprawe, 2015). In this model we will use both time and country pair fixed effects and that will allow us to have a better understanding of how corruption affects migration.

It is important to highlight that in these two models we will differentiate 12 different sub-samples according to three layers. The first one is overall migration and corresponds to the full dataset. This layer has been studied in most studies of this type, *e.g.* Poprawe (2015). The second layer is the differentiation according to the skill level, that is, high, medium and low skill migration subsamples. Usually only high skilled migration is studied, for example Dimant *et al.* (2013). The third layer allows us to differentiate these four subsamples (overall and high, medium and low skilled) by gender (male and female). This layer has not been studied in any other work as far as we know.

Notice there are other possible models in the literature that could also be used for the purpose of a robustness check (Negative Binomial, Zero-inflated Poisson (ZIP)) and Burger *et al.* (2009) suggest that these different specifications should all be presented. We will present these alternative models regressions in appendix.

4.1. Poisson regression with robust standard errors (PPML)

Starting with the Poisson regression with robust standard errors for 2010 that is Table 6 (and 7 presents the same information but through incidence rate ratios), we can see that in all regressions the pseudo R-squared is fairly high varying between 0.65 and 0.80, indicating a good fit for our model (Dimant *et al.* (2013). R-squared is less than 0.30, and even though these two cannot be compared it can still be used as a reference. In addition, our results suggest that our variables are better at explaining migration for the high skilled. The Wald test also does not seem to point to some kind of trouble.

Furthermore, we can see that overall (for all migrants) the results point in the same direction predicted by the literature. The interpretation of the coefficients is as follows: a one point change in the CPI score of origin will lead to an expected increase of migration by a factor of ($e^{0.233}$) 1.262. In fact, the CPI (Corruption Perception Index) scores of the origin and destination countries are significant for the migration movement, in the sense that considering all migrants the effect of the CPI score of the origin country has a significant negative impact on overall migration stocks, and that the CPI score of the destination country has a significant positive coefficient.

It is important to remember that a higher CPI score corresponds to a country with less perceived corruption and that a lower score corresponds to higher perceived corruption in a certain country. Therefore, according to our results, origin countries with higher perceived corruption, *ceteris paribus*, tend to have a higher number of migrants reflecting a push effect caused by corruption. This result is in accordance with Ahmad and Arjumand (2016) and their claim that corruption may drive those who do not want to comply with it to leave their homelands.

On the other hand, destinations countries with a lower perceived corruption are more attractive to migrants. This pull effect can be explained by the view that less corruption is associated with lower levels of uncertainty and smaller monetary costs associated with corruption (Ketterer and Rodríguez-Pose, 2015). These arrival countries thus provide better conditions than their higher corruption counterparts for migrants that, *ceteris paribus*, will rather migrate into these countries with lower levels of corruption.

These results can be compared with the results previously stated in the literature. There is a slight disagreement in the literature since, for instance, Ariu and Squicciarini (2013) suggest that corruption acts more as a deterrent for inflows than as a catalyser for outflows. While on the other hand, Poprawe (2015) finds that the effect of corruption in the origin country (push effect) is stronger than the pull effect of the (lack of) perceived corruption in the destination.

We find that even though the CPI score of the destination country has a higher absolute value for overall migration, validating that the pull effect of the arrival country is stronger than the push of the home country although not by much, it is still less significant than the CPI score of the origin country. So, our results do not fully support either side.

Notice however that compared to these studies we have a rather homogenous and small set of destination countries with low perceived corruption levels that may be diminishing the effect of the CPI score on the destination country.

Our results are similar to the above empirical works and to the study by Dimant *et al.* (2013) that found that higher levels of corruption in the origin country lead to more outward migration and less inward migration, also matching theoretical models such as the one presented by Mariani (2007). So, there is some consensus in the literature that more corruption leads to less attractive conditions for individuals even though there is a discussion on the scale of these effects and on which one is stronger.

An extra layer of analysis allows us to differentiate this effect by skill level. We find that the push effect of the CPI score of the origin country is stronger for migrants with lower skill level in a seemingly linear relationship, that is the lower the skill level the stronger will be the push effect reflected on migration.[7] This goes against Dimant *et al.* (2013) result that the negative effect of corruption is stronger for higher skilled migrants, since they have made a higher investment in their skill level and want to retrieve their investment. It seems that higher levels of skill are able to handle corruption better than lower levels, it could be that higher skilled workers can either be more efficient at protecting themselves or their possessions (*e.g.* by using part of their income as suggested by Mariani (2007)). Higher skilled individuals may also have an easier access to corruption networks, lessening their losses or even providing them some rents (Swamy *et al.*, 2001) since corruption, as defined by the CPI, implies taking advantage of a position of power and it is reasonable to assume that higher skilled workers have better chances to have such jobs. Even if we assume that all skill levels are equally likely to have such a position of power, incentives to become corrupt may be different since, for instance, if the punishment for being found using a certain position to obtain gains is being fired, it would be easier for the high skilled individual to find a replacement job and so, *ceteris paribus*, he/she would also be more likely to try to engage in such activities.

If we go even further it is possible to assume that better connected individuals have a stronger chance of becoming high skilled than other individuals, taking advantage of their favourable position in a way similar to what is described by Yusuf (2012).

However, it is possible to reconcile our results with Dimant *et al.* (2013), since the CPI score of the destination country has a significant effect on high skilled migrants, and both effects combined seem to generate a stronger overall effect on high skilled migrants, so there is still a strong incentive for these individuals to migrate since, despite the previous

[7] See Table 4 for skill related differences in coefficients.

reasoning, other countries may still be able to provide better opportunities for these individuals. Unlike what happens for the high skilled migrants, where the effect of the country of arrival is stronger than that for the home country, for medium skill migrants these two effects have similar strength. Low skilled migrants do not seem to be affected by the CPI score of the destination country, so the strength of the effects of the CPI score of the home country by skill is opposite to the one of the destination country, that is the effect of the CPI score of the destination country is stronger for migrants with a high skill level and for the origin country the effect is stronger for lower skilled migrants.

If we also differentiate between genders, we can see that both genders present similar effects of corruption (in both destination and origin) for migration (without discriminating by skill level).[8] However, the same does not happen when we do discriminate the results by gender and skill. When it comes to high skilled migrants there is significantly higher coefficients for female migrants for the CPI score of the destination and for the origin CPI score of the destination country. It appears that for high skilled female migrants there might be an additional need to prevent against corruption and they might be more likely to be targeted/discriminated and face difficulties that their male counterparts do not face (Docquier *et al.*, 2009). Other reasons presented by the literature are that females may be less able to enter these corruption networks (Swamy *et al.*, 2001), or that they might have a lower tolerance against corruption compared to male high skilled individuals (Jha and Sarangi, 2018). This results in a stronger pull effect for female migrants and therefore at this point our suspicion of an additional incentive to migrate seems to be accurate.

Furthermore, the lack of a significant difference between coefficients for the CPI score of the destination country for high skilled men and women is also a very interesting effect. It is possible that given that our destination countries are developed countries that grant women and men equal rights, the pull effect is similar between both genders (the difference between coefficients is non-significant). On the other hand, if similar rights are not granted then the negative repercussion of targeting a woman may be lower and therefore there is an incentive to target these individuals (Shleifer and Vishny, 1993). Taking this into account, high skilled female individuals may not be *au pair* with their male counterparts and either have more losses (or lack of gains) due to corruption or, for instance, may have to spend more resources in protection. In either case there is an additional incentive for this

[8] See Table 5 for gender related differences in coefficients.

segment of the population to migrate, which is consistent with the results found for the CPI score of the origin country.

However, this effect may not hold for every skill segment. In fact, there is a significantly higher effect of the CPI score of the origin country for low skilled male migrants than for the same skill level females and this may be due to the role of the man as the income provider of the family. Seeing corruption as a kind of taxation on their income reducing it below sustainable values, these individuals might have the main responsibility to go abroad in order to find income for their families (Docquier *et al.*, 2009) and the previous effect may not be as stronger as we thought.

A counter argument to this may be that, given the low income we expect low skilled individuals to have, then low skilled females are more sensitive to corruption than males in the sense that after accounting for the costs of corruption they might not be able to support the costs of migration. In this scenario, the extra loss of value corruption implies for women not only reduces their income, just like for male low skilled migrants, but may also turn the option of migrating less viable for females. Therefore, the main difference between gender differences for low and high skilled migrants is that high skilled females (assuming a corresponding higher level of income) can afford to move despite the negative influence of corruption and have an incentive therefore to do so, whereas low skilled may not be able to. This may hide (bias) the gender differences (reflected in the regression coefficients) that exist in a country.

As for medium skilled female migrants, it seems that they are not as restricted in their ability to move as low skilled females and, on the other hand, they have a lower incentive to migrate than high skilled migrants since their gains will be more limited. Therefore, the differences between female and male in this segment are not as pronounced and therefore non-significant.

As for the CPI of the destination country, the coefficients are not significant for low skilled migrants. These results may be underestimated due to selectiveness of immigration laws in OECD countries that gives an advantage to higher skilled migrants. The coefficients are also not significant for medium skilled female migrants suggesting a lower selectivity based on the corruption criteria in these cases. Furthermore, since destination countries (OECD countries) already have a high control of corruption (the average CPI score for the destination countries in 2010 is 7.965 whereas the average CPI score for the origin countries is around 4, as we can see in Table 3), there may not be a big gain difference due to less corruption for lower skilled migrants between these destination countries.

The same reasoning may be applied for the CPI score of the destination country for female medium skilled migrants. However, for male medium skilled migrants this effect is significant, which may be due to the gender role previously mentioned where the lower uncertainty or costs caused by low corruption affect male medium migrants since they are the ones who are more likely to migrate.

We are aware that our results might change with an enlargement in the sample of destination countries besides OECD countries, including those with low CPI score, since the difference in possible gains of migration would be different between countries in such a situation, immigration laws might be less restrictive and that is something we must bear in mind.

In conclusion, our results support that for higher skilled migrants the strongest effect is the pull effect of (lack of) corruption, that is, the CPI score of the destination country has a higher absolute coefficient value than the one for the origin country. On the other hand, for lower skilled migrants the strongest of these (and most significant) effects is the push effect caused by corruption at the origin country. It seems that high skilled migrants may have a higher ability to select the countries they migrate to as stated in Grogger and Hanson (2011), seem to face a lower pressure due to corruption in origin and may be better at prospering in such an environment. On the other hand, lower skilled migrants seem to be more likely to be forced to move out due to adverse conditions (Ivlevs and King, 2017) rather than being attracted by external motivators.

As for the gender dimension of the migration phenomenon, we also find interesting results. First, we find that corruption at the home country is stronger for female high skilled migrants, reflecting an additional incentive to migrate as predicted by Docquier *et al.* (2009). We also point out some reasons why gender differences between lower skilled migration (low and medium skill) may not exhibit a similar behaviour than the previously described. As for corruption at the destination country, gender differences are not as pronounced since for high skilled these differences are not significant and for low skilled neither the male or female coefficient is actually significant. For medium skilled migrants this effect is more relevant to male migrants possibly due to the gender role of the male as an income provider. Tables 4 and 5 resume our findings.

Table 4 - Differences in coefficients by skill level

Variable	Skill Levels Compared	Difference by skill is significant?	Wald test (equality of coefficients)	Higher effect
CPI_Origin	Low-Medium	Yes	11.29 (0.0008)	Low
	Medium-High	Yes	5.83 (0.0157)	Medium
	High-Low	Yes	18.42 (0.0000)	Low
CPI_Destination	Low-Medium	No	0.68 (0.4104)	-
	Medium-High	Yes	10.21 (0.0014)	High
	High-Low	Yes	9.06 (0.0026)	High

Source: Own computation.

Table 5 - Differences in coefficients by skill level and gender

Variable	Skill Level	Difference by gender is significant?	Wald test (equality of coefficients)	Higher effect
CPI_Origin	All	No	0.04 (0.8377)	-
	Low	Yes	4.50 (0.0338)	Male
	Medium	No	0.08 (0.7720)	-
	High	Yes	6.13 (0.0133)	Female
CPI_Destination	All	No	0.01 (0.9275)	-
	Low	Yes	4.09 (0.0432)	Female
	Medium	Yes	18.36 (0.0000)	Male
	High	No	0.09 (0.7653)	-

Source: Own computation; Note: P-value in parenthesis.

Table 6- Poisson Regression with Robust Standard Errors (PPML) - 2010

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.233**	0.117	0.207**	0.506***	0.238**	0.174	0.110	0.517***	0.235***	0.068	0.292***	0.501***
	(0.091)	(0.123)	(0.095)	(0.114)	(0.099)	(0.131)	(0.103)	(0.124)	(0.090)	(0.122)	(0.098)	(0.110)
CPI_Origin	-0.265***	-0.429***	-0.261***	-0.140***	-0.268***	-0.407***	-0.257***	-0.165***	-0.263***	-0.451***	-0.268***	-0.115**
	(0.073)	(0.084)	(0.080)	(0.053)	(0.069)	(0.082)	(0.076)	(0.054)	(0.078)	(0.087)	(0.087)	(0.054)
Colonial_Relationship	1.193***	1.988***	0.703***	0.996***	1.131***	1.855***	0.581**	0.982***	1.260***	2.145***	0.835***	1.011***
	(0.208)	(0.359)	(0.240)	(0.194)	(0.186)	(0.339)	(0.227)	(0.235)	(0.246)	(0.383)	(0.272)	(0.165)
Common_Language	0.512**	-0.152	0.406	0.974***	0.609***	-0.030	0.571**	1.019***	0.412*	-0.293	0.241	0.927***
	(0.221)	(0.262)	(0.253)	(0.165)	(0.206)	(0.255)	(0.225)	(0.165)	(0.244)	(0.275)	(0.297)	(0.173)
Common_Religion	0.714**	0.881	0.823**	0.361	0.821***	0.898*	0.985***	0.531*	0.594*	0.866	0.625*	0.177
	(0.322)	(0.550)	(0.347)	(0.296)	(0.318)	(0.529)	(0.343)	(0.309)	(0.339)	(0.585)	(0.368)	(0.295)
Contiguity	0.793	1.050**	0.646	0.289	0.763*	1.077**	0.631	0.248	0.818	1.006**	0.652	0.332
	(0.488)	(0.438)	(0.474)	(0.399)	(0.459)	(0.441)	(0.436)	(0.371)	(0.518)	(0.436)	(0.516)	(0.435)
Distance	-0.625***	-1.013***	-0.723***	-0.370***	-0.571***	-0.914***	-0.649***	-0.358***	-0.683***	-1.124***	-0.798***	-0.384***
	(0.112)	(0.141)	(0.133)	(0.096)	(0.112)	(0.144)	(0.132)	(0.100)	(0.115)	(0.143)	(0.140)	(0.095)
GDPPC_Destination	0.707**	0.250	1.045***	0.884***	1.185***	0.650*	1.827***	1.201***	0.347	-0.104	0.648	0.576**
	(0.291)	(0.399)	(0.336)	(0.222)	(0.261)	(0.354)	(0.286)	(0.213)	(0.332)	(0.471)	(0.394)	(0.249)
GDPPC_Origin	0.307***	0.278***	0.291***	0.329***	0.337***	0.291***	0.317***	0.365***	0.279***	0.264**	0.274***	0.291***
	(0.087)	(0.107)	(0.098)	(0.082)	(0.086)	(0.106)	(0.099)	(0.082)	(0.090)	(0.110)	(0.101)	(0.084)
Inflation_Destination	0.952***	0.468*	1.449***	1.142***	0.920***	0.430*	1.280***	1.192***	0.979***	0.494**	1.595***	1.086***
	(0.207)	(0.239)	(0.248)	(0.176)	(0.194)	(0.230)	(0.223)	(0.176)	(0.224)	(0.250)	(0.285)	(0.185)
Inflation_Origin	-0.040*	-0.062	-0.053**	-0.015	-0.040*	-0.056	-0.048*	-0.022	-0.041*	-0.068*	-0.057**	-0.008
	(0.023)	(0.038)	(0.027)	(0.022)	(0.023)	(0.038)	(0.025)	(0.024)	(0.024)	(0.040)	(0.029)	(0.021)
Unemp_Destination	-0.043	-0.149**	0.030	-0.010	-0.014	-0.115**	0.063	0.009	-0.062	-0.180***	0.024	-0.029
	(0.049)	(0.061)	(0.058)	(0.041)	(0.048)	(0.058)	(0.057)	(0.042)	(0.054)	(0.066)	(0.063)	(0.044)
Unemp_Origin	-0.031**	-0.038**	-0.019	-0.036***	-0.034**	-0.034*	-0.026	-0.039***	-0.029*	-0.043**	-0.014	-0.032**
	(0.014)	(0.019)	(0.016)	(0.013)	(0.014)	(0.019)	(0.017)	(0.013)	(0.015)	(0.020)	(0.017)	(0.013)
Pop_Destination	0.837***	1.035***	0.674***	0.796***	0.864***	1.047***	0.728***	0.814***	0.824***	1.040***	0.646***	0.784***
	(0.063)	(0.089)	(0.059)	(0.057)	(0.062)	(0.086)	(0.058)	(0.061)	(0.067)	(0.097)	(0.063)	(0.058)
Pop_Origin	0.558***	0.533***	0.511***	0.593***	0.548***	0.529***	0.506***	0.578***	0.568***	0.539***	0.514***	0.609***
	(0.033)	(0.050)	(0.035)	(0.031)	(0.032)	(0.049)	(0.035)	(0.031)	(0.035)	(0.053)	(0.037)	(0.033)
Tax_Destination	-0.074***	-0.036*	-0.122***	-0.095***	-0.062***	-0.027	-0.101***	-0.088***	-0.084***	-0.045**	-0.138***	-0.102***
	(0.013)	(0.021)	(0.014)	(0.012)	(0.012)	(0.019)	(0.013)	(0.012)	(0.014)	(0.022)	(0.016)	(0.012)
Tax_Origin	0.019	0.014	0.011	0.020*	0.016	0.015	0.007	0.016	0.022*	0.014	0.015	0.024**
	(0.012)	(0.013)	(0.014)	(0.012)	(0.013)	(0.013)	(0.015)	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
pseudo R^2	0.730	0.717	0.685	0.770	0.731	0.707	0.698	0.760	0.718	0.720	0.662	0.769
Wald chi2(17)	1307.878	458.601	904.908	1928.317	1582.268	486.623	900.278	2001.426	1013.328	423.445	782.228	1945.902
Prob > chi2	0	0	0	0	0	0	0	0	0	0	0	0
Log pseudolikelihood	-4.176e+07	-1.825e+07	-1.394e+07	-1.325e+07	-2.090e+07	-9260434.188	-6561661.590	-7224430.601	-2.211e+07	-9375783.447	-7921172.832	-6485112.832

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Table 7- Poisson Regression with Robust Standard Errors (PPML) 2010 - Incidence Rate Ratio

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	1.262**	1.125	1.230**	1.658***	1.269**	1.190	1.116	1.676***	1.264***	1.070	1.339***	1.651***
	(0.115)	(0.138)	(0.117)	(0.189)	(0.126)	(0.156)	(0.115)	(0.209)	(0.113)	(0.130)	(0.131)	(0.182)
CPI_Origin	0.767***	0.651***	0.770***	0.869***	0.765***	0.665***	0.774***	0.848***	0.769***	0.637***	0.765***	0.892**
	(0.056)	(0.055)	(0.062)	(0.046)	(0.053)	(0.055)	(0.059)	(0.046)	(0.060)	(0.055)	(0.066)	(0.049)
Colonial_Relationship	3.297***	7.297***	2.020***	2.707***	3.100***	6.392***	1.788**	2.670***	3.526***	8.538***	2.305***	2.748***
	(0.686)	(2.620)	(0.486)	(0.524)	(0.578)	(2.169)	(0.406)	(0.626)	(0.867)	(3.269)	(0.626)	(0.452)
Common_Language	1.669**	0.859	1.501	2.647***	1.839***	0.971	1.770**	2.772***	1.510*	0.746	1.272	2.526***
	(0.369)	(0.225)	(0.379)	(0.437)	(0.378)	(0.248)	(0.398)	(0.456)	(0.368)	(0.205)	(0.378)	(0.436)
Common_Religion	2.042**	2.412	2.277**	1.435	2.273***	2.455*	2.677***	1.701*	1.811*	2.377	1.868*	1.193
	(0.657)	(1.326)	(0.791)	(0.424)	(0.722)	(1.299)	(0.917)	(0.526)	(0.615)	(1.391)	(0.687)	(0.352)
Contiguity	2.211	2.857**	1.908	1.335	2.144*	2.935**	1.879	1.282	2.266	2.734**	1.920	1.393
	(1.080)	(1.250)	(0.904)	(0.532)	(0.984)	(1.294)	(0.819)	(0.475)	(1.174)	(1.191)	(0.990)	(0.606)
Distance	0.535***	0.363***	0.485***	0.691***	0.565***	0.401***	0.522***	0.699***	0.505***	0.325***	0.450***	0.681***
	(0.060)	(0.051)	(0.065)	(0.066)	(0.063)	(0.058)	(0.069)	(0.070)	(0.058)	(0.046)	(0.063)	(0.065)
GDPPC_Destination	2.028**	1.284	2.843***	2.420***	3.271***	1.916*	6.214***	3.322***	1.415	0.902	1.913	1.779**
	(0.590)	(0.512)	(0.955)	(0.536)	(0.852)	(0.679)	(1.774)	(0.708)	(0.470)	(0.424)	(0.754)	(0.443)
GDPPC_Origin	1.360***	1.321***	1.338***	1.389***	1.400***	1.338***	1.374***	1.440***	1.321***	1.303**	1.315***	1.338***
	(0.118)	(0.141)	(0.132)	(0.114)	(0.120)	(0.142)	(0.136)	(0.118)	(0.119)	(0.143)	(0.133)	(0.113)
Inflation_Destination	2.592***	1.596*	4.261***	3.133***	2.508***	1.537*	3.596***	3.294***	2.661***	1.638**	4.930***	2.963***
	(0.535)	(0.381)	(1.056)	(0.552)	(0.486)	(0.353)	(0.803)	(0.579)	(0.595)	(0.410)	(1.407)	(0.547)
Inflation_Origin	0.960*	0.940	0.949**	0.985	0.960*	0.945	0.953*	0.978	0.960*	0.934*	0.944**	0.992
	(0.022)	(0.036)	(0.025)	(0.022)	(0.022)	(0.036)	(0.024)	(0.023)	(0.023)	(0.037)	(0.027)	(0.021)
Unemp_Destination	0.958	0.861**	1.031	0.990	0.986	0.892**	1.065	1.009	0.940	0.835***	1.024	0.972
	(0.047)	(0.052)	(0.060)	(0.041)	(0.047)	(0.052)	(0.060)	(0.042)	(0.050)	(0.055)	(0.065)	(0.043)
Unemp_Origin	0.969**	0.963**	0.981	0.965***	0.967**	0.967*	0.974	0.961***	0.972*	0.958**	0.986	0.968**
	(0.014)	(0.018)	(0.016)	(0.013)	(0.014)	(0.018)	(0.016)	(0.013)	(0.014)	(0.019)	(0.016)	(0.012)
Pop_Destination	2.309***	2.816***	1.962***	2.217***	2.372***	2.849***	2.071***	2.258***	2.280***	2.829***	1.907***	2.191***
	(0.145)	(0.251)	(0.115)	(0.127)	(0.146)	(0.245)	(0.121)	(0.137)	(0.153)	(0.275)	(0.120)	(0.126)
Pop_Origin	1.747***	1.705***	1.666***	1.810***	1.731***	1.697***	1.659***	1.783***	1.765***	1.715***	1.672***	1.838***
	(0.057)	(0.086)	(0.058)	(0.056)	(0.055)	(0.083)	(0.059)	(0.055)	(0.062)	(0.091)	(0.063)	(0.060)
Tax_Destination	0.929***	0.964*	0.885***	0.910***	0.940***	0.974	0.904***	0.916***	0.920***	0.956**	0.871***	0.903***
	(0.012)	(0.020)	(0.013)	(0.011)	(0.011)	(0.019)	(0.012)	(0.011)	(0.013)	(0.021)	(0.014)	(0.011)
Tax_Origin	1.019	1.014	1.011	1.020*	1.016	1.015	1.007	1.017	1.022*	1.014	1.015	1.024**
	(0.013)	(0.013)	(0.014)	(0.012)	(0.013)	(0.013)	(0.015)	(0.012)	(0.013)	(0.013)	(0.014)	(0.012)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
pseudo R^2	0.730	0.717	0.685	0.770	0.731	0.707	0.698	0.760	0.718	0.720	0.662	0.769
Wald chi2(17)	1307.878	458.601	904.908	1928.317	1582.268	486.623	900.278	2001.426	1013.328	423.445	782.228	1945.902
Prob > chi2	0	0	0	0	0	0	0	0	0	0	0	0
Log pseudolikelihood	-4.176e+07	-1.825e+07	-1.394e+07	-1.325e+07	-2.090e+07	-9260434.188	-6561661.590	-7224430.601	-2.211e+07	-9375783.447	-7921172.832	-6485112.832

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Exponentiated Coefficients; Source: Own Computation.

Control variables

Regarding control variables, as we can see in Table 6 (and in Table 7), our results are also very close to what is shown in the literature. Starting by the existence of a colonial relationship between the pairs of countries, we can see that not only this variable is significant, but it also has a large positive coefficient in line with the literature, for example Docquier *et al.* (2007).

All coefficients of the different subsamples are significant at the 1% level, except for female medium skilled migrants where the coefficient is significant at 5%. This difference between male and female medium skilled migrants may be due to the already mentioned different gender roles, that lead to this difference since males are more likely to migrate to provide income for their families.

This gender difference isn't as clear for low skilled migrants since they depend more on the selectivity of immigration laws that give an advantage to higher skilled individual and have a lower ability to adapt in the destination country (for instance, they may be able to find employment more easily). Therefore, if a colonial relationship lowers migration restrictive requirements then the pull/push effect would be stronger for all lower skilled individuals and we would not be able to see these gender differences.

It is easy to see why a colonial a relationship can provide positive and large coefficients, since many colonies share similarities in their laws and also some cultural traits that improve the adaptation to the destination country (Beine *et al.*, 2016) Furthermore, as previously implied, many of these destinations countries provide some benefits or less obstacles (less restrictive immigration laws) to the migration of individuals from related countries due to their relationship. If two countries share a Colonial Relationship, *ceteris paribus*, then, on average, overall migration between both countries is 3.297 ($=e^{1.193}$) more likely than if they do not share such a trait. This interpretation is similar for other dummy variables and for different subsamples.

Other cultural variables such as common language and common religion dummies have positive and significant results at a 5% level for overall migration, as expected by Beine *et al.* (2016), although Mayda (2010) finds common language to be non-significant. This implies that countries that share these traits will be more likely to have more migration between them.

As stated, sharing a common religion also has a significant and positive effect, however this effect is not equal among different genders or skill levels. In fact, for male migrants subsamples the significance of this coefficient is always lower than for their female counterparts, suggesting that a common religion is more likely to influence the decision to migrate of female individuals. Furthermore, for both male and female migrants, the medium skilled *stratum* is the one where a common religion is the most significant.

When controlling for (not only but especially) colonial relationship we can see that having a common language is significant for overall migration. While it is surprising to see this lack of effect on lower skilled migrants (possibly since colonial relationship partially overlaps the common language dummy), the same cannot be said about high skilled migrants whose result is positive as we expected, possibly because a higher skilled migrant has the ability to select the best fitting country for him, whereas lower skilled migrants might move to where they can and not to where they wish and therefore may not be able to make decisions based on common language.

The results of the contiguity dummy seem to point also in the same direction since, despite being non-significant for overall migration (this result is also obtained by Mayda (2010), this variable is only significant on the low skill sample (higher coefficient for females), and these migrants are the most likely to migrate into neighbour countries possibly due to the increased difficulties and costs that they might not be able to support in more distant (geographical and culturally since neighbour countries share some cultural traits) countries.

A different explanation suggests that having a common official primary language does not mean that both populations speak the same language. For example, we can point out the country pairs (origin-destination respectively) of Angola and Portugal or of Ivory Coast and France. In fact, although they share a common language, the origin countries have other national languages. It is reasonable to assume that lower levels of education will have less contact with the Portuguese/French language and will be less likely to actually know how to speak these languages, and rather use a different local language for their everyday life (*e.g.* creole languages). On the other hand, higher skilled individuals may learn this language since they have more contact with it in their education years and may be more likely to use such a language on an everyday basis. Therefore, this result would demonstrate precisely this point, namely that high skilled individuals, unlike low skilled, are influenced by a common language because they do know and use the official common language.

Within this specification of the model, the GDP per capita of the destination country performs as expected for overall migrations with a positive coefficient. For instance, for overall migration this coefficient has the interpretation of an elasticity, that is a change in

1% of GDP per capita turns on average, *ceteris paribus*, migration 0.707% more likely. This result is also obtained in the literature (*e.g.* Poprawe, 2015). Theoretically we expect wealthier countries to attract more migrants (Beine *et al.*, 2016), and this result holds in most regressions. In fact, the effect of the GDP per capita in the destination country is significant for medium and high skilled migrants, unlike low skilled probably because they are more likely to be limited on their ability to move out of their origin country (they are less likely to be able to afford it or to compete for visas). Furthermore, income in these countries might not be significantly different for low skilled individuals and so their gains would be rather stable among these countries and this would be reflected on the lack of significance.

This result may also be explained by Vogler and Rotte (2000)'s suggestion that relative income might also explain this result since individuals might have a higher aversion for being poor when surrounded by wealthier individuals than when surrounded by a poor community. This may happen since a lower skill level is also likely to be reflected in a lower income; this might also help explaining our results since it is likely that higher skilled individuals are better remunerated for their skills in more developed countries, thereby increasing relative poverty of these migrants and reducing the incentives to migrate. This relative poorness aversion also seems stronger for male migrants than for female for low skilled migration.

If we also differentiate this result by gender, we see that this effect is always significant and with positive coefficients for female migrants, but it is only significant for high skilled males and even in this case it is less significant than the female counterpart. These results may also be linked to a gender pay gap, where countries with lower differences between genders attract female migrants. These results suggest that discrimination through income may be a cause of concern for female migrants even for these countries with high development.

This reasoning seems to support our previous hypothesis that discrimination of gender may have two different sources. First, discrimination reflects itself as an income differential between male and female migrants and in this case, there are still significant differences even between destination countries. Second, discrimination might result in an incentive for corruption to target female individuals as previously explained and, in this case, destination countries do not seem to have significant difference between them.

As for the main gravity variables the results are as expected. Namely, the population destination countries have a pull effect on migrants and the population of origin countries with higher population have a push effect. Furthermore, the coefficients are higher for the pull than for the push effect, a result similar to the one obtained by Poprawe (2015). There are no differences of significance between male and female migrants, nor between different skill levels. Furthermore, as distance increases there is a dissuasion effect on migration, reflecting higher costs associated with dislocation as predicted, for example, in Dimant *et al.* (2013). These might be monetary costs or psychological ones as Vogler and Rotte (2000) mention. Similarly, the effect for the population variables is also equally significant among subsamples.

As for the unemployment variable the results are quite interesting. The results point in the sense that the unemployment rate at the destination country has no significant effect in overall migration, unlike what is predicted by White and Buehler (2018). However, their selection of destination countries is also larger and that may explain the difference between these results. Furthermore, this variable is only significant for low skilled migrants reinforcing our assumption that they are less able to compete in the labour market and therefore are more sensitive to this variable. It also has a stronger effect on low skilled male migrants than on female migrants, reinforcing their role as income providers for lower skilled (or educated) families and this seems also to be reflected on the corresponding origin variable for low skilled migrants.

The unemployment rate at origin has a negative coefficient for overall migration, and it can be interpreted as: a 1 percentage point increase in the unemployment rate at origin, *ceteris paribus*, on average will decrease migration by a factor of 0.969 (1-0.969=3.1% reduction of migration). Two effects may be causing this disturbance from expectations, that is higher unemployment at origin should lead to more migration (Beine *et al.*, 2016). The first one is that the lack of employment might mean lack of income and therefore migrants may not be able to afford moving out of the country. If we also consider the effect of the GDP per capita of the origin country, we see that higher levels of this variable lead to more migration suggesting that the lack of income might be a barrier to migration and that, as this becomes less of a burden, individuals will be then more likely to migrate. This effect is similar in significance across subsamples.

The effect of the level of unemployment at origin might be amplified since even though individuals might be unemployed they might still receive some social benefits in their home countries that they might lose, as suggested by Vogler and Rotte (2000).

If we differentiate among different skill levels, it seems that this explanation suits better the low skilled individuals. As for high skilled, the high significance may be explained by what they have to lose in case of an unsuccessful migration (uncertainty cost.) and a similar case may be made for high skilled females *vs.* their male counterparts.

Moving forward, even though our taxation variable is a rather broad proxy for taxation, the results seem to fit the literature when point that the taxation rate is not significant for migration at origin (except for high skilled males). On the other hand, taxation at destination does negatively impact the decision to migrate. These results are very similar to those obtained by Poprawe (2015). It is easy to understand that, all else constant, an individual would rather pay less taxes and have more available income especially for higher skilled since the income that is being taxed is higher. This effect is lower for low skilled migrants since they are expected.to have a lower income and therefore pay lower taxes. Furthermore, it is possible that, since these individuals are the most likely to depend on certain social benefits provided by the destination, a lower taxation may not be as important due to the effect of this trade off (Beine *et al.*, 2016). It is also at this segment (low skill) that there are differences in the significance level between both genders, which relate to our previous claim of different gender roles among genders.

Finally, we also consider the inflation rate just like Poprawe (2015). We find out that inflation at the home country decreases migration between a given country pair, and the coefficient of this variable increases with the skill level of the migrant leading, and overall it has the same effect for both genders. The results suggest high skilled are better able to support the negative effects of inflation, namely the loss of real income that is implied (on the long run also leads to a weakening of the currency).

As for the destination country the coefficients of the inflation rate are positive, large and significant. Although a higher inflation might suggest a higher macroeconomic instability, we suspect that due to the small amount of destination countries and to the fact that in 2010 many of them shared a common currency (the euro), these results may be misleading. In fact, the inflation rate in these countries, in 2010, varies between 0.89% and 3.6%, that are rather stable levels. So, we must take into consideration that all destination countries are in a relatively stable environment and that this effect would be less pronounced if we had more years into the analysis. This effect is rather stable among both genders and by skill level.

We will now present the results for the model with two-way fixed effects (time and country pair) and compare them with what we found and with what is already established in the literature, to see if our results hold or if there may be some additional effects.

4.2. Fixed effects regression

In this section, we extend the previous analysis by estimating a panel data model that controls for fixed effects for country pairs and also for years.[9] It is important to notice that fixed effects models have been used as a sort of robustness checks in the literature.[10] This happens for instance in Dimant *et al.* (2013) and in Poprawe (2015). In a nutshell, we can say that our results (see Table 8) diverge from the ones showed in the previous section, but the conclusion that corruption is an important factor on the decision to migrate is maintained.

Note that, according to Greene (2011), controlling for country pairs fixed effects eliminates some bias that may arise from omitted variables that do not change across time but that do change over country pair, for example distance or common official language (it does not change in this period) between countries; it is now accounted in this term. On the other hand, controlling for time fixed effects eliminates the bias that may arise from variables that affect all the country pairs, for instance, the existence of a global financial crisis that affects the results for 2010. The interpretation of coefficients is similar to the previous model.

In this model, our analysis takes into account both the variety of CPI scores of destination and origin countries and the changes in this score across time. The results show an inversion of the sign of the coefficient for the CPI score at the origin country.

However, this result can still be explained by the theoretical mechanisms previously suggested. In fact, if we consider that in order to migrate an individual must have the ability to cover the implied costs (a certain threshold for income that when reached makes the individual move into another country), then an increase in the CPI score implies a decrease in perceived corruption and therefore a lower cost for individuals (a view of corruption as a kind of taxation on income). This lower expenditure will bring the income value of individuals closer to the migration threshold and therefore lead to higher migration stocks between a given country pair. In a different perspective, this effect of an increase of

---

[9] Note that the test for the relevance of time fixed effects is provided in appendix 1.

[10] As previously mentioned, the usage of different specifications has also been suggested as a robustness test. So, even though we will not analyze the negative binomial and the zero inflated Poisson regressions, we will present these regressions in the appendix 2 to 5.

corruption across time, that implies lower migration stocks from the same origin country to a destination, may be stronger than the effect of a similar change of the corruption level across countries, that implies higher migration stocks between country pairs and therefore lead to this change in the expected sign of the coefficient.

The significance of coefficients is similar across gender and across different skill levels. However, if we discriminate our results for both gender and skill level some differences arise. The significance level is lower for low skilled as they are further apart this migration threshold, and so this effect is lessened in this subsample since they may not be able to afford the costs of migration, even if they want to. Furthermore, despite overall migration does not show a difference in significance levels, for each single skill level the effect is more significant for females than for males, so the gender differences that arise due to corruption in the origin are more pronounced in this model than in the last.

As for the CPI score of the destination country, it generally behaves in a similar way to the one that is proposed in the previous section, although there is no effect of corruption in the destination country on high skilled migration, in opposition to what occurs in the previous model. However, if high skilled migrants do have a better ability to handle corruption then this result might just mean that, given the destination countries, there is not a significant difference on corruption levels that justifies migration. In fact, Table 9 shows that the CPI score of destination country loses its attraction power on high skilled migrants faster than, for example, medium skilled. Furthermore, the maximum attraction point is reached at a CPI score of 7.835 a value that is lower than many of our destinations countries scores (89 cases out of 140 possible). As for the difference in significance between medium and low skilled individuals, this may be due to a higher proximity of the migration decision threshold (low skilled are less likely to be able to afford dislocations).

As for gender differences, they only emerge in the low skilled segment, exposing the effect we had claimed to be hidden, that is corruption may take an extra value from female migrants that further distances them from this threshold. Table 9 supports this finding and further approaches the panel and regular Poisson models. In fact, gender role effects and the fact that female migrants may not be able to afford migration also appear to be reflected in these regressions. Furthermore, the reverse role in medium skilled migration reflect the gains from the reduction of gender related corruption and that this segment of the population can afford migration costs. Finally, high skilled female migrants have a less significant reduction of the attraction of this score, also reflecting an extra cost of corruption.

Table 8 - Poisson two-way fixed effects regression with robust standard errors, 1980-2010

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.140***	0.128**	0.172***	0.0392	0.133***	0.121**	0.185***	0.0298	0.146***	0.130*	0.155***	0.0494
	(0.0338)	(0.0591)	(0.0421)	(0.0365)	(0.0343)	(0.0564)	(0.0448)	(0.0384)	(0.0368)	(0.0667)	(0.0442)	(0.0381)
CPI_Origin	0.0495**	0.0753**	0.0970**	0.0470**	0.0519**	0.0773**	0.109***	0.0553***	0.0477**	0.0724*	0.0864**	0.0412**
	(0.0223)	(0.0361)	(0.0385)	(0.0200)	(0.0221)	(0.0328)	(0.0369)	(0.0204)	(0.0237)	(0.0438)	(0.0398)	(0.0209)
GDPPC_Destination	0.468***	0.464***	0.363***	0.546***	0.413***	0.341**	0.302**	0.493***	0.527***	0.611***	0.419***	0.586***
	(0.102)	(0.171)	(0.126)	(0.0838)	(0.108)	(0.151)	(0.141)	(0.0851)	(0.104)	(0.217)	(0.125)	(0.0864)
GDPPC_Origin	0.234***	0.157*	0.327***	0.237***	0.269***	0.197**	0.373***	0.268***	0.196**	0.110	0.279**	0.207***
	(0.0802)	(0.0941)	(0.102)	(0.0667)	(0.0780)	(0.0848)	(0.0980)	(0.0650)	(0.0846)	(0.111)	(0.109)	(0.0712)
Inflation_Destination	-0.0248***	-0.0213**	-0.0268**	-0.0440***	-0.0158	-0.00904	-0.000303	-0.0557***	-0.0273***	-0.0254**	-0.0349***	-0.0367***
	(0.00838)	(0.0106)	(0.0120)	(0.0104)	(0.0130)	(0.0140)	(0.0199)	(0.0139)	(0.00791)	(0.0107)	(0.0114)	(0.00861)
Inflation_Origin	-0.000288***	-0.000184	-0.000474***	-0.000170**	-0.000329***	-0.000232**	-0.000468***	-0.000227***	-0.000240***	-0.000131	-0.000460***	-0.000104
	(0.0000755)	(0.000123)	(0.0000964)	(0.0000726)	(0.0000683)	(0.0000924)	(0.0000977)	(0.0000663)	(0.0000896)	(0.000184)	(0.000113)	(0.0000772)
Unemp_Destination	0.00292	0.0373**	-0.00930	-0.0228**	0.00235	0.0344**	0.000283	-0.0323***	0.00402	0.0409**	-0.0175*	-0.0140
	(0.00967)	(0.0158)	(0.0103)	(0.00977)	(0.0103)	(0.0148)	(0.0127)	(0.0106)	(0.0103)	(0.0183)	(0.0106)	(0.00986)
Unemp_Origin	0.0148***	0.0121*	0.0253**	0.0108**	0.0161***	0.0123*	0.0271***	0.0144***	0.0133**	0.0115*	0.0229**	0.00792
	(0.00564)	(0.00658)	(0.0100)	(0.00469)	(0.00554)	(0.00693)	(0.00981)	(0.00472)	(0.00591)	(0.00665)	(0.0102)	(0.00493)
Pop_Destination	1.346**	0.781	2.086**	-1.365***	0.924	0.327	1.289	-1.686***	1.677**	1.121	2.450**	-0.956**
	(0.661)	(1.034)	(0.937)	(0.457)	(0.614)	(0.951)	(0.894)	(0.477)	(0.725)	(1.131)	(1.003)	(0.473)
Pop_Origin	2.326***	2.870***	2.212***	1.614***	2.259***	2.744***	2.181***	1.526***	2.409***	3.013***	2.273***	1.720***
	(0.359)	(0.581)	(0.584)	(0.238)	(0.329)	(0.532)	(0.562)	(0.249)	(0.395)	(0.626)	(0.601)	(0.241)
Tax_Destination	-0.0248***	-0.0192	-0.0130	-0.0100	-0.0188*	-0.00982	-0.00739	-0.00628	-0.0320***	-0.0299*	-0.0225**	-0.0152
	(0.00923)	(0.0154)	(0.00848)	(0.00893)	(0.00962)	(0.0156)	(0.00870)	(0.00936)	(0.00941)	(0.0160)	(0.00914)	(0.00935)
Tax_Origin	-0.00233	0.00161	0.00299	0.0107*	-0.00326	0.00378	0.00441	0.00929	-0.00107	-0.000491	0.00217	0.0129**
	(0.00645)	(0.00846)	(0.00911)	(0.00555)	(0.00654)	(0.00812)	(0.00889)	(0.00591)	(0.00663)	(0.00934)	(0.00963)	(0.00555)
N	6478	6415	6418	6465	6372	6248	6259	6286	6453	6322	6333	6400

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Table 9 - Poisson two-way fixed effects regression with robust standard errors (with squared CPI of destination country), 1980-2010

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.439**	0.175	0.750***	0.514***	0.401**	0.246	0.657**	0.433***	0.463**	0.839***	0.0858	0.647***
	(0.188)	(0.275)	(0.254)	(0.164)	(0.200)	(0.291)	(0.258)	(0.164)	(0.208)	(0.285)	(0.310)	(0.176)
$CPI\_Destination^2$	-0.0207	-0.00321	-0.0424**	-0.0328***	-0.0185	-0.00843	-0.0343*	-0.0280**	-0.0220	-0.0503**	0.00300	-0.0412***
	(0.0126)	(0.0184)	(0.0187)	(0.0112)	(0.0132)	(0.0195)	(0.0184)	(0.0111)	(0.0140)	(0.0209)	(0.0206)	(0.0122)
CPI_Origin	0.0496**	0.0754**	0.0992***	0.0475**	0.0524**	0.0775**	0.112***	0.0559***	0.0474**	0.0878**	0.0724*	0.0413**
	(0.0218)	(0.0360)	(0.0379)	(0.0198)	(0.0217)	(0.0324)	(0.0366)	(0.0203)	(0.0232)	(0.0388)	(0.0440)	(0.0206)
GDPPC_Destination	0.492***	0.467***	0.432***	0.581***	0.446***	0.353**	0.392**	0.532***	0.543***	0.474***	0.609***	0.619***
	(0.105)	(0.178)	(0.129)	(0.0848)	(0.115)	(0.163)	(0.153)	(0.0861)	(0.108)	(0.129)	(0.220)	(0.0875)
GDPPC_Origin	0.237***	0.158*	0.330***	0.240***	0.271***	0.198**	0.376***	0.271***	0.198**	0.282***	0.109	0.210***
	(0.0798)	(0.0945)	(0.102)	(0.0660)	(0.0777)	(0.0851)	(0.0979)	(0.0644)	(0.0842)	(0.109)	(0.112)	(0.0703)
Inflation_Destination	-0.0241***	-0.0211**	-0.0265**	-0.0454***	-0.0153	-0.00842	-0.000216	-0.0589***	-0.0266***	-0.0342***	-0.0255**	-0.0370***
	(0.00820)	(0.0102)	(0.0118)	(0.0104)	(0.0127)	(0.0130)	(0.0197)	(0.0140)	(0.00790)	(0.0114)	(0.0106)	(0.00844)
Inflation_Origin	-0.000280***	-0.000182	-0.000458***	-0.000160**	-0.000323***	-0.000228**	-0.000458***	-0.000219***	-0.000230**	-0.000438***	-0.000133	-0.0000901
	(0.0000765)	(0.000124)	(0.0000955)	(0.0000734)	(0.0000683)	(0.0000942)	(0.0000964)	(0.0000663)	(0.0000916)	(0.000114)	(0.000182)	(0.0000790)
Unemp_Destination	0.00703	0.0379**	-0.000802	-0.0163*	0.00673	0.0361**	0.00940	-0.0264***	0.00779	-0.00901	0.0405**	-0.00647
	(0.00998)	(0.0159)	(0.0114)	(0.00947)	(0.0108)	(0.0152)	(0.0144)	(0.0102)	(0.0107)	(0.0118)	(0.0184)	(0.00956)
Unemp_Origin	0.0148***	0.0121*	0.0252**	0.0107**	0.0160***	0.0123*	0.0269***	0.0142***	0.0133**	0.0229**	0.0114*	0.00774
	(0.00555)	(0.00657)	(0.00992)	(0.00466)	(0.00548)	(0.00693)	(0.00977)	(0.00469)	(0.00580)	(0.0101)	(0.00666)	(0.00489)
Pop_Destination	1.412**	0.790	2.151**	-1.264***	0.997	0.354	1.403	-1.573***	1.740**	2.505**	1.112	-0.846*
	(0.651)	(1.019)	(0.922)	(0.452)	(0.609)	(0.945)	(0.875)	(0.474)	(0.712)	(0.992)	(1.111)	(0.463)
Pop_Origin	2.321***	2.870***	2.178***	1.607***	2.253***	2.742***	2.147***	1.519***	2.407***	2.236***	3.013***	1.715***
	(0.359)	(0.581)	(0.583)	(0.238)	(0.328)	(0.532)	(0.560)	(0.249)	(0.395)	(0.600)	(0.627)	(0.241)
Tax_Destination	-0.0203**	-0.0186	-0.00167	-0.00332	-0.0146	-0.00826	0.00245	-0.000198	-0.0274***	-0.00948	-0.0305*	-0.00727
	(0.00957)	(0.0165)	(0.0108)	(0.00871)	(0.0103)	(0.0168)	(0.0113)	(0.00914)	(0.00953)	(0.0110)	(0.0169)	(0.00917)
Tax_Origin	-0.00227	0.00163	0.00339	0.0109*	-0.00322	0.00385	0.00470	0.00943	-0.000963	0.00268	-0.000523	0.0132**
	(0.00641)	(0.00849)	(0.00903)	(0.00558)	(0.00649)	(0.00812)	(0.00887)	(0.00592)	(0.00660)	(0.00953)	(0.00940)	(0.00559)
N	6478	6415	6418	6465	6372	6248	6259	6286	6453	6322	6333	6400

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Control variables

When controlling for fixed effects of time and country pairs, we can see that the variable of the population of the origin keeps the push effect (more people also means more possible migrants). This effect has a similar significance across gender and skill. As for the population at the destination country, there are significant gender differences. The population of the destination country is not significant for females, but it is for males.

Analysing the pull effect for different skill levels we can see that population at the destiny country is not significant at attracting low skilled individuals supporting the argument that, for lower levels skills, individuals are more likely to be pushed out than pulled in. This variable has a positive and significant effect for medium skilled reflecting that an increase in skill may reflect the ability to successfully migrate into these countries (Mountford, 1997).

For high skilled this effect is negative suggesting that highly populated countries attract less of these individuals. It is possible that these skills may not be as easily transferable or that their social status might be harder to maintain, also reflecting the aversion of migrants to become relatively poorer when moving as Vogler and Rotte (2000) suggest.

Our taxation and unemployment variables allow us to differentiate whether it is the pull or the push effect that has a higher effect on migrants. For instance, the unemployment at home country leads to more migration suggesting that the results obtained in 2010 were also influenced by the economic crisis and its negative effect on the economy and posterior difficulties in the financial market. The coefficients are lower for low skilled migrants since these individuals may have more difficulty at finding a replacement work that suits their skills when unemployed and, therefore, have a higher income dependence that prohibits them from migrating. The opposite effect is registered for high skilled individuals. The results are rather similar between both genders.

However, the difference between significances increases as skill level rises, culminating with a 1% significant effect for high skilled female migrants *vs.* a non-significant effect for males. It seems like male are better able to compete at origin country than female and this may be the result of some kind of discrimination against women.

On the other hand, the unemployment rate provides a significant push effect for high skilled migrants. Uncertainty of finding a job and having a successful migration increase the costs of migration and decreases this migration stocks (Beine *et al.*, 2016). For lower skilled migrants, they may be forced to move out of their country and be able to migrate despite the adverse conditions rather than choosing to move to countries with higher unemployment. Social benefits that these countries may provide may also explain the positive sign, as even unemployed these low skilled migrants may insure an income higher than the one they would obtain in their home country with these benefits (Beine *et al.*, 2016).

Concerning taxation level at origin, it seems that its effect only affects male high skilled individuals, as previously reported. It is easy to understand this since a higher income implies a higher absolute taxation (and in some countries with progressive tax rates also higher relative taxation), the high skilled migrants will pay a higher cost and therefore are also the ones that have more to gain by changing to a country where, with similar conditions, they pay less taxes and have more available income.

Taxation level at the destination country is significant and has a negative coefficient for overall migration but is not significant in any of the regressions by skill (for both genders). A deeper analysis shows that this is result is mostly for male migrants being significant in all but the high skill level. But for female migrants this is not so, and this is probably due to the role of women in migration, that is to reunite with their abroad partner although according to our results that seems to be changing. This variable loses some importance in explaining migration.

When it comes to GDP per capita, in the home country the results are very similar to the ones of the previous model, in the sense that there is a significant and positive variable that states that a country with higher income also has more overall migration. The coefficients are similar for medium and high skilled regressions. As for low skilled this effect is less significant since they may be further way from the migration threshold. In this skill level, female migrants are also more restrained by their income than their male counterparts. These results are consistent with the first model and, therefore, so is the previous justification for this effect that we presented, especially since the coefficients are very similar to the ones in our first model.

GDP per capita in the destination country has a much similar effect between skill levels and genders in this model comparing with the previous model. When taking into account the multiple year observations the effect seems to harmonize among the various *stratus* considered. So, the difficulties of having rather homogenous destination countries seem to be lessened (as expected) by the inclusion of observations for different years. Despite this, lower skilled migrants (medium and low) have a lower significance in their female subsamples.

As for the inflation rate at the home country, it seems to have a negative coefficient, similarly to what we previously described but very close to zero and female migrants seem to be more influenced by the economic stability that inflation reflects. On the other hand, the inflation rate at the destination country has a negative coefficient, as it seems that the inclusion of extra observations show that the instability related to a higher inflation (above a certain level) and the weaker currency effect (translates into a lower available income) may reduce migration. This effect is more significant for high skilled individuals (in this case gender differences are not relevant) and for male migrants reflecting a stronger pull effect caused by this factor. Table 10 presents a summary of the results we have obtained and discussed in this chapter.[11]

In the next chapter we shall state our final conclusions, as well as present some main limitations but also future research paths.

[11] We also present in appendix 6 a literature summary in order to contextualize the main literature on both migration and corruption (works that only cover one of these issues are not presented).

Table 10 – Results: a sum up

Variable	Layer[12]	Poisson – 2010	Poisson Panel 1980-2010	In Favor	Against
CPI_Destination	Overall Skill Gender Extra	Positive and Significant Sign Stronger Effect for High Skilled Migrants Constant Among Gender Females seem to be the most affected gender.	Positive and Significant Sign More Significant for Medium then Low Skilled Migrants Constant Among Gender Females seem to be the most affected gender.	Poprawe (2015) Dimant *et al.* (2013) Ariu and Squicciarini (2013)	None
CPI_Origin	Overall Skill Gender Extra	Negative and Significant Sign Effect Decreases with Skill Constant Among Gender Females seem to be the most affected gender.	Positive and Significant Sign Constant Among Skill Constant Among Gender Females seem to be the most affected gender.	Poprawe (2015) Dimant *et al.* (2013) Ariu and Squicciarini (2013)	All of those at left (Panel Model) Mariani (2007)
Colonial_Relationship	Overall Skill Gender	Positive and Significant Sign Constant Among Skill Constant Among Gender	-	Docquier *et al.* (2007) Beine *et al.* (2016) Vogler and Rotte (2000)	Mayda (2010)
Common_Language	Overall Skill Gender	Positive and Significant Sign Only Significant for High Skilled Migrants More Significative for Females	-	Beine *et al.* (2016) Vogler and Rotte (2000)	Mayda (2010)
Common_Religion	Overall Skill Gender	Positive and Significant Sign Only Significant for Medium Skilled Migrants More Significant for Females	-	Beine *et al.* (2016) Vogler and Rotte (2000)	None
Contiguity	Overall Skill Gender	Non-Significant (Positive) Only Significant for Low Skilled Migrants More Significant for Females	-	Beine *et al.* (2016) Vogler and Rotte (2000) Mayda (2010)	None
Distance	Overall Skill Gender	Positive and Significant Sign Constant Among Skill Constant Among Gender	-	Beine *et al.* (2016) Dimant *et al.* (2013) Poprawe (2015)	None
GDPPC_Destination	Overall Skill Gender Extra	Positive and Significant Sign Significant for Medium and High Skilled Migrants More Significant for Females Females seem to be the most affected gender.	Positive and Significant Sign Constant Among Skill Constant Among Gender Females seem to be the most affected gender.	Beine *et al.* (2016) Vogler and Rotte (2000) Poprawe (2015)	None
GDPPC_Origin	Overall Skill Gender Extra	Positive and Significant Sign Constant Among Skill Constant Among Gender Females seem to be the most affected gender.	Positive and Significant Sign More Significant for Medium and High Skilled Migrants More Significant for Females Females seem to be the most affected gender.	Beine *et al.* (2016) Vogler and Rotte (2000) Poprawe (2015)	None

Source Own Computation

[12] Note that: "Overall" corresponds to the regression with the same name; "Skill" refers to the three regressions contained in "Skill Level" and "Gender" corresponds to the "All Skill" regression by gender. Finally, comments about the combination of both these layers are also presented when necessary in "Extra".

Table 10 – Results: a sum up (continuation)

Variable	Layer	Poisson - 2010	Poisson Panel 1980-2010	In Favour	Against
Inflation_Destination	Overall Skill Gender	Positive and Significant Sign More Significant for Medium and High Skilled Migrants Constant Among Gender	Negative and Significant Sign More Significant for High Skilled Migrants More Significant for Males	Poprawe (2015) (In the Panel model)	Poprawe (2015) (In the 2010 model)
Inflation_Origin	Overall Skill Gender	Negative and Significant Sign Only Significant for Medium Skilled Migrants Constant Among Gender	Negative and Significant Sign More Significative for Medium then High Skilled Constant Among Gender	Poprawe (2015)	None
Unemp_Destination	Overall Skill Gender	Non-Significant (Negative) Only Significant for Low Skilled Migrants Non-Significant for Both Genders	Non-Significant (Positive) More Significative for High and Low Skilled Non-Significant for Both Genders	Beine *et al.* (2016) Mayda (2010)	White and Buehler (2018)
Unemp_Origin	Overall Skill Gender	Negative and Significant Sign More Significative for High then Low Skilled Migrants More Significant for Females	Positive and Significant Sign More Significative for Medium and High Skilled More Significant for Females	Beine *et al.* (2016)	Mayda (2010) White and Buehler (2018)
Pop_Destination	Overall Skill Gender	Positive and Significant Sign Constant Among Skill Constant Among Gender	Positive and Significant Sign Negative Sign for High and Positive for Medium Skilled More significant for Males	Beine *et al.* (2016) Vogler and Rotte (2000) Poprawe (2015)	None
Pop_Origin	Overall Skill Gender	Positive and Significant Sign Constant Among Skill Constant Among Gender	Positive and Significant Sign Constant Among Skill Constant Among Gender	Beine *et al.* (2016) Vogler and Rotte (2000) Poprawe (2015)	None
Tax_Destination	Overall Skill Gender	Negative and Significant Sign More Significative for Medium and High Skilled Migrants Constant Among Gender	Negative and Significant Sign Non-Significant across Skill Level More Significative for Males	Poprawe (2015) Grogger and Hanson (2011)	None
Tax_Origin	Overall Skill Gender	Non-Significant (Positive) Only Significant for High Skilled Migrants More Significative for Males	Non-Significant (Negative) Only Significant for High Skilled Migrants (Positive) Non-Significant for Both Genders	Poprawe (2015)	Grogger and Hanson (2011)

Source: Own Computation.

Chapter 5. Conclusions

In this dissertation, we started with a brief introduction and explanation about what motivated us into researching these complex issues of corruption and migration, namely the relevance of this work lies on the profound social and economic implications of both migration and corruption. Moreover, the significant gap in the existing literature, which we aimed to cover, also supports the relevance of our work.

To expose these gaps that we aimed to fill, we presented the literature about international migration, discussing both theoretical and empirical works. At the same time, we explored the literature on corruption in order to highlight the undesirable effects of corruption for an economy in general and for individuals in particular. We then brought studies that explored both issues to find a theoretical (and empirical) framework for our analysis.

After that we presented as clearly as possible how we planned to study this topic, namely by justifying the decisions we made related to our model and variables (like the choice of indicators). These decisions were based on finding answers for three different questions:

1. *Is corruption a significant push and/or pull factor for (overall) migrants?*
2. *Has corruption different impact between skilled and non-skilled migration?*
3. *Is this effect similar between male and female migrants?*

According to our analysis, we can answer the first research question affirmatively. Corruption is a significant push and also a significant pull factor for overall migration. We find that less perceived corruption at home increases the number of inward migrants and that a higher level of corruption at the destination countries decreases migration into these countries. This result is in accordance with previous literature. What is new about this finding is that we found this result to be significant in our panel gravity model, using not only cross-sectional data but also a time-series and this is a new dimension of analysis in the literature. However, this new dimension seems to have changed the expected signal (not only due to our regular Poisson model but also due to the literature) of the corruption at origin variable in the fixed effects model. It is important to highlight that these two results are not mutually exclusive, and that there seems to be an extra effect, namely the loss or gain of the ability to migrate by those near the migration threshold, that leads to this result.

This important result that arises through the introduction of a time dimension should be further tested and that may be a lead for further investigation. Experimenting with different datasets is necessary to see if our conclusions still hold, or if we may be mistaking this signal change with the interference of heterogenic individual effects on corruption. Some independent variables may also be tested by the usage of a different instrument. For instance, a non-perceptual corruption indicator, marginal tax rates or an indicator that measures the unemployment gap rather than the level of unemployment may provide some additional insights. Furthermore, the usage of a different government quality variable may also be a future path to follow. For instance, a political stability indicator could be used as an alternative to the CPI, like it happens in institutional literature. Finally, finding (and theoretically explaining) if there is a non-linear relationship between migration and corruption (also on gender and skill) may also be a crucial research path.

As for our second research question, throughout our work we have pointed out significant differences and given possible explanations based on our theoretical foundations about the existence of differences in the impact of corruption across skill levels. Furthermore, these differences are captured in both models, clearly supporting this finding. In summary, it seems that unlike what the literature predicted, high skilled individuals are not the most likely to migrate, but rather seem to be the less likely group to be affected by corruption. Adding to that, we found that the effect of corruption on low skilled migrants may be lessened by their inability to migrate. Another finding that is important to highlight is that corruption at the origin country affects different skilled individuals in a different way than corruption at destination.

As for the final question, in our analysis we found significant gender differences in the effect of corruption that to our knowledge have never been studied. However, if we simply compare overall female and male migration these differences are not significant. In fact, these gender differences remain hidden unless we add the extra skill level layer of analysis. When both gender and skill levels are taken into account, these gender differences emerge, and it seems to reflect an extra incentive in origin countries to target the female population segment. Our results also seem to demonstrate that this discrimination through corruption appears to be less significant in developed countries.

These extra layers of analysis, namely gender and skill, still need to be explored in this literature since there are not many studies dedicated to them. Studies with a higher level of detail may be useful; a particular case is the study of the interaction between the variables in economies with lower restrictions to migration movements such as the European Union.

The main limitation to further research is the lack of data around corruption and migration. Luckily there has been an increase in the awareness of the importance of these two issues. Another limitation may be the lack of motivation for decision makers to follow measures to end corruption and so inertia by these individuals may decrease the incentives for the study of this issue. Another limitation that can also be seen as an opportunity is that both these issues are always evolving. So, it is possibly that due to lack of access to data we can only focus on explaining the past, therefore always being one step behind reality and not be able to make decisions when they are most needed.

Despite this, we believe our findings are still important for decision makers linked to issues of economic development in general, and migration in particular. The existence of higher levels corruption in lower developed countries may prove to harm these economies in the short and long run. Corruption has a repelling effect for individuals, driving away important human resources and harming more than just the simple wealth of individuals, affecting their well-being. Since lower skilled individuals are also more likely to have lower income, this effect is even more hurtful since they may not have the possibility to move out or protect themselves. Therefore, our results point out the need to control corruption, especially in the cases where individuals cannot protect themselves.

Finally, our results also point to the need of protecting female individuals. A higher effort to promote equality may prove useful to decline an excessive brain drain. This is particularly important given that the education level of females is particularly important for economic growth (Docquier *et al.*, 2009).

References

Acemoglu, D., Johnson, S. and Robinson, J. A. (2002), "Reversal of fortune: Geography and institutions in the making of the modern world income distribution", *Quarterly Journal of Economics*, Vol. 117, Nº 4, pp. 1231-1294.

Aggrey, N. (2012), "Determinants of corruption at district levels in Uganda", *International Journal of Economic Perspectives*, Vol. 6, Nº 4, pp. 67-72.

Ahmad, N. and Arjumand, S. (2016), "Impact of corruption on GDP per capita through international migration: an empirical investigation", *Quality and Quantity*, Vol. 50, Nº 4, pp. 1633-1643.

Aidt, T. S. (2009), "Corruption, institutions, and economic development", *Oxford Review of Economic Policy*, Vol. 25, Nº 2, pp. 271-291.

Al-Marhubi, F. A. (2000), "Corruption and inflation", *Economics Letters*, Vol. 66, Nº 2, pp. 199-202.

Allison, P. D. and Waterman, R. P. (2002), "Fixed-effects negative binomial regression models", *Sociological Methodology*, Vol. 32, Nº, pp. 247-265.

Anderson, J. E. and Marcouiller, D. (2002), "Insecurity and the pattern of trade: An empirical investigation", *Review of Economics and Statistics*, Vol. 84, Nº 2, pp. 342-352.

Ariu, A. and Squicciarini, M. P. (2013), "The balance of brains - Corruption and migration: Corruption could be an important factor for emigration and immigration decisions by highly skilled professionals", *EMBO Reports*, Vol. 14, Nº 6, pp. 502-504.

Baltagi, B., Egger, P. and Pfaffermayr, M. (2014) 'Panel Data Gravity Models of International Trade', CESifo Group Munich.

Bardhan, P. (1997), "Corruption and Development: A Review of Issues", *Journal of Economic Literature*, Vol. 35, Nº 3, pp. 1320-1346.

Beine, M., Bertoli, S. and Fernández-Huertas Moraga, J. (2016), "A Practitioners' Guide to Gravity Models of International Migration", *World Economy*, Vol. 39, Nº 4, pp. 496-512.

Bénassy-Quéré, A., Coupet, M. and Mayer, T. (2007), "Institutional determinants of foreign direct investment", *World Economy*, Vol. 30, Nº 5, pp. 764-782.

Blackburn, K. and Powell, J. (2011), "Corruption, inflation and growth", *Economics Letters*, Vol. 113, Nº 3, pp. 225-227.

Branisa, B., Klasen, S. and Ziegler, M. (2013), "Gender Inequality in Social Institutions and Gendered Development Outcomes", *World Development*, Vol. 45, Nº, pp. 252-268.

Brücker, H., Capuano, S. and Marfouk, A. (2013), "Education, gender and international migration: insights from a panel-dataset 1980-2010", *mimeo*, Institute for Employment Research, Nuremberg

Burger, M., van Oort, F. and Linders, G. J. (2009), "On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation", *Spatial Economic Analysis*, Vol. 4, Nº 2, pp. 167-190.

Cameron, A. and Trivedi, P. (2009) *Microeconometrics Using Stata*. StataCorp LP.

Dimant, E., Krieger, T. and Meierrieks, D. (2013), "The effect of corruption on migration,1985-2000", *Applied Economics Letters*, Vol. 20, Nº 13, pp. 1270-1274.

Docquier, F., Lohest, O. and Marfouk, A. (2007), "Brain drain in developing countries", *World Bank Economic Review*, Vol. 21, Nº 2, pp. 193-218.

Docquier, F., Lowell, B. L. and Marfouk, A. (2009), "A gendered assessment of highly skilled emigration", *Population and Development Review*, Vol. 35, Nº 2, pp. 297-321.

Eicher, T., García-Pẽalosa, C. and Ypersele, T. (2009), "Education, corruption, and the distribution of income", *Journal of Economic Growth*, Vol. 14, Nº 3, pp. 205-231.

Fernández-Val, I. and Weidner, M. (2016), "Individual and time effects in nonlinear panel models with large N, T", *Journal of Econometrics*, Vol. 192, Nº 1, pp. 291-312.

Fisman, R. and Svensson, J. (2007), "Are corruption and taxation really harmful to growth? Firm level evidence", *Journal of Development Economics*, Vol. 83, Nº 1, pp. 63-75.

Greene, W. H. (2011) *Econometric Analysis*. 7th edn. New York: Prentice Hall.

Grogger, J. and Hanson, G. H. (2011), "Income maximization and the selection and sorting of international migrants", *Journal of Development Economics*, Vol. 95, Nº 1, pp. 42-57.

Guimarães, P. (2008), "The fixed effects negative binomial model revisited", *Economics Letters*, Vol. 99, Nº 1, pp. 63-66.

Gupta, S., Davoodi, H. and Alonso-Terme, R. (2002), "Does corruption affect income inequality and poverty?", *Economics of Governance*, Vol. 3, Nº 1, pp. 23-45.

Heywood, P. M. and Rose, J. (2013), "Close but no Cigar: The measurement of corruption", *Journal of Public Policy*, Vol. 34, Nº 3, pp. 507-529.

Hilbe, J. (2014) *Modeling Count Data*. Cambridge University Press.

Issifou, I. (2017), "Can migration reduce civil conflicts as an antidote to rent-seeking?", *Economic Systems*, Vol. 41, Nº 3, pp. 333-353.

Ivlevs, A. and King, R. M. (2017), "Does emigration reduce corruption?", *Public Choice*, Vol. 171, Nº 3-4, pp. 389-408.

Jain, A. K. (2001), "Corruption: A review", *Journal of Economic Surveys*, Vol. 15, Nº 1, pp. 71-116.

Jha, C. K. and Sarangi, S. (2018), "Women and corruption: What positions must they hold to make a difference?", *Journal of Economic Behavior & Organization*, Vol. 151, Nº, pp. 219-233.

Ketterer, T. D. and Rodríguez-Pose, A. (2015), "Local quality of government and voting with one's feet", *Annals of Regional Science*, Vol. 55, Nº 2-3, pp. 501-532.

Maria Davidescu, A. A., Strat, V. A., Grosu, R. M. and Zgură, I. D. (2017), "Determinants Of Romanians' migration within the European Union: Static and dynamic panel gravity approaches", *Amfiteatru Economic*, Vol. 19, Nº 46, pp. 621-639.

Mariani, F. (2007), "Migration as an antidote to rent-seeking?", *Journal of Development Economics*, Vol. 84, Nº 2, pp. 609-630.

Mauro, P. (1995), "Corruption and growth", *Quarterly Journal of Economics*, Vol. 110, Nº 3, pp. 681-712.

Mayda, A. M. (2010), "International migration: A panel data analysis of the determinants of bilateral flows", *Journal of Population Economics*, Vol. 23, Nº 4, pp. 1249-1274.

Mountford, A. (1997), "Can a brain drain be good for growth in the source economy?", *Journal of Development Economics*, Vol. 53, Nº 2, pp. 287-303.

Özden, C., Parsons, C. R., Schiff, M. and Walmsley, T. L. (2011), "Where on earth is everybody? The evolution of global bilateral migration 1960-2000", *World Bank Economic Review*, Vol. 25, Nº 1, pp. 12-56.

Peng, B. (2009), "Rent-seeking activities and the 'brain gain' effects of migration", *Canadian Journal of Economics*, Vol. 42, Nº 4, pp. 1561-1577.

Poprawe, M. (2015), "On the relationship between corruption and migration: empirical evidence from a gravity model of migration", *Public Choice*, Vol. 163, Nº 3-4, pp. 337-354.

Rodrik, D., Subramanian, A. and Trebbi, F. (2004), "Institutions rule: The primacy of institutions over geography and integration in economic development", *Journal of Economic Growth*, Vol. 9, Nº 2, pp. 131-165.

Santos Silva, J. M. C. and Tenreyro, S. (2006), "The log of gravity", *Review of Economics and Statistics*, Vol. 88, Nº 4, pp. 641-658.

Shleifer, A. and Vishny, R. W. (1993), "Corruption", *Quarterly Journal of Economics*, Vol. 108, Nº 3, pp. 599-617.

Steinberg, D. (2017), "Resource shocks and human capital stocks – Brain drain or brain gain?", *Journal of Development Economics*, Vol. 127, Nº, pp. 250-268.

Swamy, A., Knack, S., Lee, Y. and Azfar, O. (2001), "Gender and corruption", *Journal of Development Economics*, Vol. 64, Nº 1, pp. 25-55.

Tarek, B. A. and Ahmed, Z. (2017), "Institutional Quality and Public Debt Accumulation: An Empirical Analysis", *International Economic Journal*, Vol. 31, Nº 3, pp. 415-435.

Vogler, M. and Rotte, R. (2000), "The effects of development on migration: Theoretical issues and new empirical evidence", *Journal of Population Economics*, Vol. 13, Nº 3, pp. 485-508.

White, R. and Buehler, D. (2018), "A closer look at the determinants of international migration: decomposing cultural distance", *Applied Economics*, Vol. 50, Nº 33, pp. 3575-3595.

Wooldridge, J. (1999), "Distribution-free estimation of some nonlinear panel data models", *Journal of Econometrics*, Vol. 90, Nº 1, pp. 77-97.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*. 2nd edn. The MIT Press.

Yusuf, H. O. (2012), "'High value' migration and complicity in underdevelopment and corruption in the global South: Receiving from the attic", *Third World Quarterly*, Vol. 33, Nº 3, pp. 441-457.

Appendices

Appendix 1 - Year fixed effects redundancy test

Gender-Skill Level	Wald Test	P-value
Overall_Migrants	85.37	0.0000
Low Skill_Migrants	43.34	0.0000
Medium Skill_Migrants	87.07	0.0000
High Skill_Migrants	104.11	0.0000
Male_Migrants	75.45	0.0000
Male_Low Skill_Migrants	41.96	0.0000
Male_Medium Skill_Migrants	89.58	0.0000
Male_High Skill_Migrants	88.17	0.0000
Female_ Migrants	61.06	0.0000
Female_Low Skill_Migrants	36.91	0.0000
Female_Medium Skill_Migrants	59.45	0.0000
Female_High Skill_Migrants	115.33	0.0000

Source: Own computation

Appendix 2 - Negative Binomial regression with robust standard errors, 2010

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.412***	0.358***	0.397***	0.459***	0.290***	0.273***	0.241***	0.355***	0.467***	0.369***	0.463***	0.536***
	(0.075)	(0.087)	(0.079)	(0.071)	(0.073)	(0.087)	(0.077)	(0.070)	(0.080)	(0.090)	(0.084)	(0.076)
CPI_Origin	-0.068	-0.177***	-0.035	-0.048	-0.097**	-0.197***	-0.070	-0.074*	-0.045	-0.155**	-0.012	-0.028
	(0.046)	(0.060)	(0.048)	(0.039)	(0.046)	(0.059)	(0.048)	(0.041)	(0.048)	(0.062)	(0.051)	(0.040)
Colonial_Relationship	2.447***	2.858***	2.066***	1.921***	2.674***	3.156***	2.432***	1.969***	2.372***	2.780***	1.935***	1.927***
	(0.406)	(0.409)	(0.539)	(0.373)	(0.482)	(0.499)	(0.647)	(0.433)	(0.378)	(0.372)	(0.495)	(0.338)
Common_Language	1.094***	0.752***	1.150***	1.476***	1.308***	0.989***	1.449***	1.588***	1.025***	0.647***	1.060***	1.424***
	(0.155)	(0.229)	(0.158)	(0.131)	(0.165)	(0.235)	(0.178)	(0.142)	(0.156)	(0.228)	(0.162)	(0.129)
Common_Religion	0.902***	1.593***	0.451**	0.051	0.779***	1.376***	0.345	0.086	0.881***	1.624***	0.380	-0.020
	(0.284)	(0.353)	(0.224)	(0.195)	(0.275)	(0.339)	(0.225)	(0.193)	(0.296)	(0.363)	(0.240)	(0.207)
Contiguity	-0.344	-0.064	-0.326	-0.465*	-0.134	0.256	-0.105	-0.402	-0.542*	-0.305	-0.512*	-0.514**
	(0.274)	(0.350)	(0.262)	(0.258)	(0.298)	(0.369)	(0.286)	(0.299)	(0.282)	(0.389)	(0.272)	(0.243)
Distance	-0.879***	-0.924***	-1.000***	-0.773***	-0.807***	-0.838***	-0.897***	-0.741***	-0.956***	-1.019***	-1.087***	-0.806***
	(0.074)	(0.091)	(0.072)	(0.066)	(0.077)	(0.092)	(0.075)	(0.072)	(0.078)	(0.096)	(0.076)	(0.067)
GDPPC_Destination	0.614***	0.556**	0.207	0.745***	1.433***	1.528***	1.152***	1.230***	0.277	0.171	-0.111	0.448**
	(0.209)	(0.246)	(0.178)	(0.170)	(0.223)	(0.262)	(0.201)	(0.173)	(0.222)	(0.259)	(0.197)	(0.189)
GDPPC_Origin	0.340***	0.266**	0.331***	0.508***	0.409***	0.327***	0.412***	0.572***	0.299***	0.238**	0.288***	0.458***
	(0.082)	(0.108)	(0.078)	(0.068)	(0.085)	(0.111)	(0.083)	(0.074)	(0.083)	(0.108)	(0.080)	(0.067)
Inflation_Destination	0.997***	0.718***	1.021***	1.024***	0.875***	0.634***	0.795***	0.972***	1.047***	0.737***	1.128***	1.036***
	(0.126)	(0.149)	(0.119)	(0.112)	(0.123)	(0.144)	(0.123)	(0.115)	(0.136)	(0.161)	(0.132)	(0.121)
Inflation_Origin	-0.022	-0.049***	-0.015	0.011	-0.024	-0.051***	-0.018	0.013	-0.018	-0.039**	-0.010	0.009
	(0.016)	(0.018)	(0.017)	(0.016)	(0.017)	(0.018)	(0.017)	(0.017)	(0.016)	(0.020)	(0.017)	(0.016)
Unemp_Destination	0.008	-0.056	0.028	0.021	0.001	-0.064	0.036	0.013	0.010	-0.054	0.030	0.022
	(0.039)	(0.045)	(0.039)	(0.035)	(0.039)	(0.044)	(0.040)	(0.035)	(0.042)	(0.049)	(0.041)	(0.038)
Unemp_Origin	0.036***	0.028**	0.038***	0.035***	0.035***	0.028**	0.034***	0.035***	0.036***	0.024**	0.041***	0.035***
	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)	(0.011)	(0.011)	(0.012)	(0.010)
Pop_Destination	1.041***	1.028***	1.053***	1.172***	1.147***	1.167***	1.143***	1.247***	0.988***	0.954***	1.022***	1.113***
	(0.063)	(0.069)	(0.054)	(0.043)	(0.066)	(0.071)	(0.059)	(0.049)	(0.065)	(0.073)	(0.057)	(0.044)
Pop_Origin	0.752***	0.706***	0.777***	0.794***	0.748***	0.695***	0.778***	0.785***	0.753***	0.710***	0.767***	0.796***
	(0.034)	(0.045)	(0.032)	(0.029)	(0.033)	(0.043)	(0.032)	(0.030)	(0.037)	(0.047)	(0.034)	(0.031)
Tax_Destination	-0.053***	-0.038***	-0.040***	-0.054***	-0.032***	-0.017	-0.017	-0.039***	-0.063***	-0.045***	-0.046***	-0.066***
	(0.011)	(0.013)	(0.010)	(0.009)	(0.011)	(0.013)	(0.011)	(0.010)	(0.012)	(0.015)	(0.011)	(0.009)
Tax_Origin	0.020**	0.042***	0.014	0.006	0.016*	0.034***	0.010	0.003	0.019**	0.043***	0.014	0.005
	(0.009)	(0.011)	(0.009)	(0.008)	(0.009)	(0.011)	(0.009)	(0.008)	(0.009)	(0.011)	(0.010)	(0.008)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
pseudo R^2	0.056	0.059	0.060	0.071	0.058	0.062	0.064	0.071	0.059	0.063	0.063	0.076
Wald chi2(17)	1569.982	997.107	2032.094	2671.563	1697.389	1065.431	2044.447	2659.335	1405.346	920.503	1818.795	2480.436
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Log pseudolikelihood	-19852.412	-16840.853	-17011.165	-17618.794	-18046.610	-15049.902	-15122.270	-15899.866	-18252.906	-15159.116	-15441.278	-16001.937
Alpha	2.662	3.113	2.822	2.365	2.878	3.372	3.070	2.676	2.714	3.255	2.955	2.387

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Appendix 3 - Negative Binomial regression with robust standard errors, 2010 – Incidence Rate Ratio

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	-1.147***	-0.958***	-1.125***	-1.138***	-0.703***	-0.599***	-0.689***	-0.701***	-1.121***	-0.941***	-1.088***	-1.078***
	(0.118)	(0.098)	(0.105)	(0.111)	(0.091)	(0.083)	(0.085)	(0.090)	(0.115)	(0.093)	(0.098)	(0.106)
CPI_Origin	-0.009	-0.037	-0.017	-0.058	-0.071	-0.045	-0.048	-0.086	-0.058	-0.016	-0.075	-0.121*
	(0.078)	(0.068)	(0.073)	(0.077)	(0.072)	(0.063)	(0.067)	(0.070)	(0.074)	(0.064)	(0.069)	(0.072)
Colonial_Relationship	-20.223***	-21.610***	-20.866***	-20.253***	-20.845***	-21.290***	-20.478***	-1.372	-20.545***	-22.209***	-21.112***	-21.877***
	(0.494)	(0.455)	(0.474)	(0.523)	(0.583)	(0.449)	(0.570)	(1.127)	(0.492)	(0.452)	(0.467)	(0.568)
Common_Language	-1.315***	-0.897***	-0.946***	-1.327***	-1.672***	-0.904***	-1.247***	-1.827***	-1.166***	-0.281	-0.851***	-1.008***
	(0.375)	(0.294)	(0.288)	(0.340)	(0.346)	(0.252)	(0.270)	(0.319)	(0.349)	(0.246)	(0.272)	(0.312)
Common_Religion	0.076	-0.218	-0.297	-0.079	0.298	-0.233	-0.025	-0.069	0.031	-0.185	-0.013	-0.049
	(0.364)	(0.342)	(0.346)	(0.363)	(0.344)	(0.326)	(0.329)	(0.348)	(0.350)	(0.319)	(0.317)	(0.346)
Contiguity	1.498*	1.747**	1.608*	1.864**	0.955	0.840	0.770	1.220	1.643*	1.438*	1.347	1.798*
	(0.898)	(0.884)	(0.903)	(0.942)	(0.900)	(0.923)	(0.937)	(0.954)	(0.927)	(0.859)	(0.885)	(0.961)
Distance	0.936***	1.094***	1.038***	1.083***	0.858***	1.014***	0.944***	1.062***	0.997***	1.129***	1.029***	1.105***
	(0.147)	(0.147)	(0.141)	(0.154)	(0.140)	(0.133)	(0.132)	(0.149)	(0.146)	(0.139)	(0.130)	(0.148)
GDPPC_Destination	0.363	0.216	0.437*	0.307	-0.863***	-1.046***	-0.944***	-1.239***	0.275	-0.022	0.436*	0.270
	(0.238)	(0.225)	(0.236)	(0.238)	(0.217)	(0.205)	(0.211)	(0.216)	(0.232)	(0.210)	(0.234)	(0.238)
GDPPC_Origin	-0.549***	-0.484***	-0.621***	-0.551***	-0.567***	-0.576***	-0.690***	-0.656***	-0.519***	-0.514***	-0.573***	-0.503***
	(0.125)	(0.111)	(0.116)	(0.122)	(0.111)	(0.100)	(0.103)	(0.106)	(0.117)	(0.102)	(0.107)	(0.112)
Inflation_Destination	-0.604***	-0.405***	-0.557***	-0.538***	0.082	0.180	0.194	0.166	-0.625***	-0.460***	-0.443***	-0.490***
	(0.157)	(0.146)	(0.155)	(0.155)	(0.151)	(0.140)	(0.151)	(0.152)	(0.151)	(0.137)	(0.151)	(0.152)
Inflation_Origin	-0.052*	-0.051**	-0.057**	-0.039	-0.033	-0.044**	-0.033	-0.050**	-0.056**	-0.041*	-0.066***	-0.023
	(0.027)	(0.024)	(0.024)	(0.025)	(0.024)	(0.022)	(0.022)	(0.023)	(0.025)	(0.022)	(0.023)	(0.024)
Unemp_Destination	-0.063	-0.083	-0.120*	-0.063	0.007	-0.055	-0.074	-0.055	-0.057	-0.125**	-0.160**	-0.027
	(0.080)	(0.065)	(0.068)	(0.076)	(0.072)	(0.059)	(0.060)	(0.067)	(0.076)	(0.060)	(0.063)	(0.070)
Unemp_Origin	0.005	0.006	0.000	-0.002	-0.000	-0.001	-0.011	-0.017	-0.006	0.015	-0.005	-0.010
	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)	(0.013)	(0.011)	(0.012)	(0.012)
Pop_Destination	-0.508***	-0.411***	-0.550***	-0.558***	-0.712***	-0.621***	-0.799***	-0.779***	-0.530***	-0.409***	-0.631***	-0.646***
	(0.099)	(0.087)	(0.092)	(0.100)	(0.098)	(0.083)	(0.091)	(0.098)	(0.097)	(0.080)	(0.090)	(0.098)
Pop_Origin	-0.691***	-0.739***	-0.773***	-0.799***	-0.706***	-0.759***	-0.778***	-0.825***	-0.740***	-0.681***	-0.783***	-0.874***
	(0.061)	(0.055)	(0.059)	(0.062)	(0.056)	(0.052)	(0.054)	(0.058)	(0.062)	(0.049)	(0.056)	(0.064)
Tax_Destination	0.080***	0.057***	0.051***	0.068***	0.019	-0.006	-0.026	-0.018	0.077***	0.045***	0.029*	0.078***
	(0.020)	(0.018)	(0.017)	(0.019)	(0.017)	(0.016)	(0.016)	(0.017)	(0.019)	(0.016)	(0.016)	(0.018)
Tax_Origin	-0.023*	-0.020	-0.015	-0.025*	-0.012	-0.009	-0.003	-0.009	-0.011	-0.008	-0.017	-0.010
	(0.014)	(0.012)	(0.013)	(0.013)	(0.013)	(0.011)	(0.012)	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
pseudo R^2	0.056	0.059	0.060	0.071	0.058	0.062	0.064	0.071	0.059	0.063	0.063	0.076
Wald chi2(17)	1569.982	997.107	2032.094	2671.563	1697.389	1065.431	2044.447	2659.335	1405.346	920.503	1818.795	2480.436
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Log pseudolikelihood	-19852.412	-16840.853	-17011.165	-17618.794	-18046.610	-15049.902	-15122.270	-15899.866	-18252.906	-15159.116	-15441.278	-16001.937
Alpha	2.662	3.113	2.822	2.365	2.878	3.372	3.070	2.676	2.714	3.255	2.955	2.387

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Exponentiated Coefficients; Source: Own Computation.

Appendix 4 – Zero Inflated Poisson regression with robust standard errors, 2010 – Logit (1)

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.318***	0.383***	0.325***	0.321***	0.495***	0.550***	0.502***	0.496***	0.326***	0.390***	0.337***	0.340***
	(0.037)	(0.038)	(0.034)	(0.036)	(0.045)	(0.046)	(0.043)	(0.045)	(0.038)	(0.036)	(0.033)	(0.036)
CPI_Origin	0.991	0.964	0.983	0.944	0.932	0.956	0.953	0.918	0.944	0.984	0.927	0.886*
	(0.077)	(0.066)	(0.072)	(0.072)	(0.067)	(0.060)	(0.064)	(0.064)	(0.070)	(0.063)	(0.064)	(0.064)
Colonial_Relationship	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.254	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.286)	(0.000)	(0.000)	(0.000)	(0.000)
Common_Language	0.268***	0.408***	0.388***	0.265***	0.188***	0.405***	0.287***	0.161***	0.312***	0.755	0.427***	0.365***
	(0.101)	(0.120)	(0.112)	(0.090)	(0.065)	(0.102)	(0.078)	(0.051)	(0.109)	(0.186)	(0.116)	(0.114)
Common_Religion	1.079	0.804	0.743	0.924	1.348	0.792	0.976	0.933	1.031	0.831	0.987	0.952
	(0.393)	(0.275)	(0.257)	(0.336)	(0.464)	(0.258)	(0.321)	(0.325)	(0.361)	(0.265)	(0.313)	(0.330)
Contiguity	4.475*	5.740**	4.991*	6.452**	2.599	2.316	2.159	3.388	5.169*	4.214*	3.846	6.036*
	(4.016)	(5.074)	(4.504)	(6.076)	(2.339)	(2.138)	(2.024)	(3.232)	(4.789)	(3.621)	(3.403)	(5.800)
Distance	2.549***	2.986***	2.825***	2.954***	2.358***	2.756***	2.569***	2.892***	2.711***	3.091***	2.797***	3.018***
	(0.374)	(0.440)	(0.398)	(0.454)	(0.331)	(0.367)	(0.340)	(0.430)	(0.395)	(0.429)	(0.362)	(0.448)
GDPPC_Destination	1.438	1.241	1.547*	1.359	0.422***	0.351***	0.389***	0.290***	1.317	0.978	1.547*	1.310
	(0.342)	(0.279)	(0.365)	(0.323)	(0.092)	(0.072)	(0.082)	(0.063)	(0.306)	(0.206)	(0.362)	(0.311)
GDPPC_Origin	0.578***	0.617***	0.537***	0.576***	0.567***	0.562***	0.502***	0.519***	0.595***	0.598***	0.564***	0.605***
	(0.072)	(0.068)	(0.062)	(0.070)	(0.063)	(0.056)	(0.052)	(0.055)	(0.069)	(0.061)	(0.060)	(0.068)
Inflation_Destination	0.546***	0.667***	0.573***	0.584***	1.086	1.197	1.214	1.180	0.535***	0.631***	0.642***	0.613***
	(0.086)	(0.098)	(0.089)	(0.090)	(0.164)	(0.168)	(0.184)	(0.180)	(0.081)	(0.086)	(0.097)	(0.093)
Inflation_Origin	0.950*	0.950**	0.945**	0.961	0.968	0.957**	0.968	0.951**	0.946**	0.959*	0.936***	0.977
	(0.025)	(0.023)	(0.022)	(0.024)	(0.023)	(0.021)	(0.021)	(0.022)	(0.024)	(0.021)	(0.022)	(0.024)
Unemp_Destination	0.939	0.920	0.887*	0.939	1.007	0.946	0.928	0.947	0.945	0.882**	0.852**	0.973
	(0.075)	(0.060)	(0.060)	(0.071)	(0.072)	(0.056)	(0.056)	(0.064)	(0.071)	(0.053)	(0.054)	(0.068)
Unemp_Origin	1.005	1.006	1.000	0.998	1.000	0.999	0.989	0.984	0.994	1.015	0.995	0.990
	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)	(0.013)	(0.011)	(0.012)	(0.012)
Pop_Destination	0.602***	0.663***	0.577***	0.573***	0.491***	0.538***	0.450***	0.459***	0.588***	0.665***	0.532***	0.524***
	(0.059)	(0.058)	(0.053)	(0.058)	(0.048)	(0.045)	(0.041)	(0.045)	(0.057)	(0.053)	(0.048)	(0.052)
Pop_Origin	0.501***	0.478***	0.462***	0.450***	0.494***	0.468***	0.459***	0.438***	0.477***	0.506***	0.457***	0.417***
	(0.030)	(0.026)	(0.027)	(0.028)	(0.028)	(0.024)	(0.025)	(0.025)	(0.029)	(0.025)	(0.026)	(0.027)
Tax_Destination	1.083***	1.058***	1.052***	1.071***	1.019	0.994	0.974	0.982	1.080***	1.046***	1.030*	1.082***
	(0.021)	(0.019)	(0.018)	(0.020)	(0.017)	(0.016)	(0.015)	(0.017)	(0.020)	(0.017)	(0.017)	(0.020)
Tax_Origin	0.977*	0.981	0.985	0.975*	0.988	0.991	0.997	0.991	0.989	0.992	0.983	0.990
	(0.014)	(0.012)	(0.013)	(0.013)	(0.013)	(0.011)	(0.012)	(0.013)	(0.013)	(0.011)	(0.012)	(0.012)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
N Zeroes	191	239	243	212	252	340	336	306	211	295	292	254
Wald chi2(17)	1261.418	431.072	849.339	1956.893	1524.388	447.012	834.299	1953.643	967.371	389.504	722.711	1989.313
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Log pseudolikelihood	-4.065e+07	-1.789e+07	-1.352e+07	-1.280e+07	-2.024e+07	-9026728.188	-6305498.977	-6917219.140	-2.149e+07	-9170492.907	-7655741.537	-6253078.631

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Appendix 5 – Zero Inflated Poisson regression with robust standard errors, 2010 – Poisson (2)

Independent Variables	Overall Migrants	Skill Level			Female Migrants				Male Migrants			
		Low	Medium	High	All Skill	Low Skill	Medium Skill	High Skill	All Skill	Low Skill	Medium Skill	High Skill
CPI_Destination	0.190**	0.077	0.156	0.469***	0.191*	0.129	0.056	0.466***	0.192**	0.023	0.238**	0.462***
	(0.092)	(0.123)	(0.096)	(0.117)	(0.099)	(0.130)	(0.103)	(0.124)	(0.090)	(0.122)	(0.100)	(0.113)
CPI_Origin	-0.281***	-0.436***	-0.272***	-0.160***	-0.282***	-0.414***	-0.266***	-0.180***	-0.279***	-0.456***	-0.278***	-0.135***
	(0.071)	(0.083)	(0.078)	(0.051)	(0.067)	(0.081)	(0.074)	(0.051)	(0.077)	(0.086)	(0.086)	(0.052)
Colonial_Relationship	1.188***	1.968***	0.679***	0.987***	1.130***	1.831***	0.559**	0.979***	1.247***	2.105***	0.792***	1.001***
	(0.207)	(0.359)	(0.237)	(0.189)	(0.185)	(0.339)	(0.223)	(0.231)	(0.244)	(0.383)	(0.267)	(0.160)
Common_Language	0.512**	-0.153	0.399	0.969***	0.598***	-0.038	0.550**	0.999***	0.413*	-0.288	0.236	0.925***
	(0.214)	(0.258)	(0.245)	(0.161)	(0.199)	(0.252)	(0.218)	(0.162)	(0.235)	(0.270)	(0.288)	(0.168)
Common_Religion	0.643**	0.810	0.748**	0.290	0.732**	0.807	0.890***	0.410	0.524	0.786	0.563	0.114
	(0.320)	(0.555)	(0.341)	(0.293)	(0.316)	(0.537)	(0.336)	(0.307)	(0.337)	(0.588)	(0.360)	(0.290)
Contiguity	0.799*	1.060**	0.670	0.304	0.779*	1.103**	0.674	0.283	0.823*	1.023**	0.685	0.345
	(0.470)	(0.429)	(0.456)	(0.381)	(0.442)	(0.433)	(0.418)	(0.354)	(0.500)	(0.427)	(0.499)	(0.418)
Distance	-0.612***	-0.991***	-0.703***	-0.358***	-0.554***	-0.882***	-0.621***	-0.341***	-0.668***	-1.095***	-0.770***	-0.371***
	(0.111)	(0.140)	(0.132)	(0.095)	(0.111)	(0.142)	(0.130)	(0.099)	(0.114)	(0.141)	(0.139)	(0.094)
GDPPC_Destination	0.711**	0.254	1.049***	0.884***	1.135***	0.567	1.740***	1.097***	0.349	-0.095	0.659*	0.566**
	(0.284)	(0.390)	(0.328)	(0.219)	(0.261)	(0.365)	(0.288)	(0.217)	(0.326)	(0.461)	(0.387)	(0.248)
GDPPC_Origin	0.347***	0.303***	0.310***	0.367***	0.372***	0.311***	0.329***	0.394***	0.313***	0.281**	0.281***	0.330***
	(0.084)	(0.110)	(0.099)	(0.075)	(0.082)	(0.108)	(0.097)	(0.075)	(0.088)	(0.115)	(0.104)	(0.077)
Inflation_Destination	0.944***	0.460**	1.434***	1.133***	0.912***	0.421*	1.270***	1.183***	0.971***	0.485**	1.580***	1.078***
	(0.202)	(0.233)	(0.242)	(0.174)	(0.189)	(0.223)	(0.218)	(0.174)	(0.218)	(0.244)	(0.279)	(0.182)
Inflation_Origin	-0.034	-0.056	-0.049*	-0.010	-0.034	-0.051	-0.045*	-0.017	-0.035	-0.063	-0.057**	-0.002
	(0.022)	(0.038)	(0.026)	(0.021)	(0.022)	(0.038)	(0.024)	(0.023)	(0.023)	(0.040)	(0.029)	(0.020)
Unemp_Destination	-0.050	-0.157***	0.021	-0.018	-0.025	-0.128**	0.047	-0.006	-0.070	-0.188***	0.013	-0.038
	(0.049)	(0.060)	(0.057)	(0.041)	(0.047)	(0.057)	(0.055)	(0.042)	(0.053)	(0.065)	(0.062)	(0.044)
Unemp_Origin	-0.033**	-0.038**	-0.021	-0.038***	-0.035**	-0.033*	-0.028*	-0.041***	-0.030**	-0.043**	-0.013	-0.034***
	(0.014)	(0.019)	(0.016)	(0.013)	(0.014)	(0.019)	(0.017)	(0.013)	(0.015)	(0.020)	(0.017)	(0.013)
Pop_Destination	0.824***	1.022***	0.657***	0.787***	0.844***	1.026***	0.700***	0.792***	0.810***	1.024***	0.628***	0.773***
	(0.062)	(0.088)	(0.057)	(0.057)	(0.061)	(0.086)	(0.058)	(0.060)	(0.066)	(0.096)	(0.062)	(0.057)
Pop_Origin	0.556***	0.527***	0.500***	0.589***	0.544***	0.518***	0.491***	0.569***	0.564***	0.529***	0.498***	0.604***
	(0.033)	(0.052)	(0.036)	(0.031)	(0.032)	(0.051)	(0.036)	(0.031)	(0.035)	(0.054)	(0.038)	(0.033)
Tax_Destination	-0.075***	-0.038*	-0.123***	-0.096***	-0.064***	-0.029	-0.104***	-0.090***	-0.085***	-0.046**	-0.138***	-0.103***
	(0.013)	(0.020)	(0.014)	(0.012)	(0.012)	(0.019)	(0.013)	(0.012)	(0.014)	(0.022)	(0.016)	(0.012)
Tax_Origin	0.019	0.015	0.011	0.020*	0.016	0.015	0.007	0.017	0.022*	0.015	0.015	0.024**
	(0.012)	(0.013)	(0.014)	(0.012)	(0.013)	(0.013)	(0.015)	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)
N	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261	2261
Number of Zeroes	191	239	243	212	252	340	336	306	211	295	292	254
Wald chi2(17)	1261.418	431.072	849.339	1956.893	1524.388	447.012	834.299	1953.643	967.371	389.504	722.711	1989.313
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Log pseudolikelihood	-4.065e+07	-1.789e+07	-1.352e+07	-1.280e+07	-2.024e+07	-9026728.188	-6305498.977	-6917219.140	-2.149e+07	-9170492.907	-7655741.537	-6253078.631

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Source: Own Computation.

Appendix 6 - Literature summary

Reference	Research question (Goal)	Research method	Sample	Dependent Variables	Main Independent Variables	Main conclusion
Ariu and Squicciarini (2013)	To find if corruption could be an important factor for emigration and immigration decisions by highly skilled professionals.	Empirical - OLS	Net flows across 123 countries Between 1990 and 2000 For Migration: Docquier *et al.* (2009) For Corruption- International Country Risk Guide corruption index	Net, inflows and outflows of migrants	Corruption GDP per Capita	Highly skilled workers are mobile, flexible and have a bigger tendency to migrate. Corruption has a negative effect (decreases) on inflows and positive on outflows.
Dimant *et al.* (2013)	To examine the influence of corruption on migration	Empirical - pooled OLS and fixed effects model with Driscoll–Kraay standard errors	111 Countries Between 1985 and 2000 For Migration: Skilled migration rates For Corruption: International Country Risk Guide corruption index	Skilled Migration Rates	Corruption (+) Distance (-) Population (-) Regime Type (+) Political Instability Youth Burden Quality of Bureaucracy Trade Openness (+) Former Colony (+) GDP per Capita (- for Skilled and + For average)	Corruption is a push factor of migration, particularly for skilled migration. Corruption increases migration outflows.
Ahmad and Arjumand (2016)	To examine the impact of corruption on GDP *per capita*	Empirical - pooled OLS and fixed effects models	94 Countries From 1996 to 2010 For Migration- World Development Indicators For Corruption- CPI	Does not directly apply since migration is not the dependent variable.		Corruption may drive those who do not want to comply with it to either leave their homelands or drives them to become corrupt.
Poprawe (2015)	To show the relationship between corruption and migration	Empirical- Negative Binomial, PPML (robustness) and fixed effects model	230 Countries 2000 and 2010 (robustness check) For Migration: Özden et al. (2011) For Corruption: Corruption Perception Index	Bilateral Migration Flows Unilateral Migration Flows (robustness check)	Corruption Home (-) Corruption Destination (+) Distance (-) Population Home (+) Population Destination (+) Tax rate Destination (-) Political Stability Common Border (+) Common Language (+) Tertiary education Inflation Home (-) Inflation Destination (-) Political System Others	Countries with more corruption attract less immigrants and leads to more emigration.

Source: Own computation.

Appendix 6 - Literature summary (continuation)

Reference	Research question (Goal)	Research method	Sample	Dependent Variables	Main Independent Variables	Main conclusion
Ketterer and Rodríguez-Pose (2015)	To investigate the impact of local quality of government on regional attractiveness to migrants	Empirical - Fixed-effect instrumental variables techniques, 2SLS regressions, and an Arellano-Bond GMM	254 European Regions Between 1995 and 2009 For Migration- Eurostat Regio Database For Corruption (governance) - World Governance Indicators	Net migration flows	Corruption (+) Lagged Migration (+) Unemployment (-) Agricultural Share (-) Quality of government index (+) And others more	Less corruption is associated with lower levels of uncertainty and monetary costs.
Yusuf (2012)	To Investigate how migration laws and policies make the state potentially complicit in corruption	Theoretical Model	Does not apply.			Developed countries might be providing opportunities for certain individuals to not only move illicit funds, but also privileging and offering them additional benefits.
Steinberg (2017)	To study the relationship between resource abundance and the selectivity of migration	Theoretical model and empirical: Pooled OLS, Fixed and Random effects model. 3SLS and Dynamic Panel model	116 source and 23 destination countries. Between 1910 and 2009 For Migration- A different study	Does not apply as corruption is not directly measured.		Natural resource shocks have a positive brain drain effect in a country, which is especially relevant in countries that are more susceptible to corruption and government inefficiencies.
Issifou (2017)	To investigate whether access to migration reduces the positive effect of natural resources on the onset of civil conflicts	Empirical Logit, Fixed effects model, PPML and 2SLS.	226 Countries From 1960 to 2010 For Migration- Özden et al. (2011)	Does not apply as corruption is not directly measured.		Migration may reduce natural resource rent seeking and decrease civil conflicts.
Mariani (2007)	To develop a new mechanism through which skilled migration may influence economic performance in the sending country	Theoretical Model	Does not apply.			There should be a positive income-maximizing migration rate. Also, inefficient allocation of talent is the reason of underdevelopment.
Peng (2009)	To study migration and rent-seeking activities in a framework of heterogeneous ability.	Theoretical Model	Does not apply.			Given the possibility of migration that increases the productive sector attractiveness, such will result not only in a quantitative movement to the productive sector but also in a qualitative movement.

Source: Own computation.